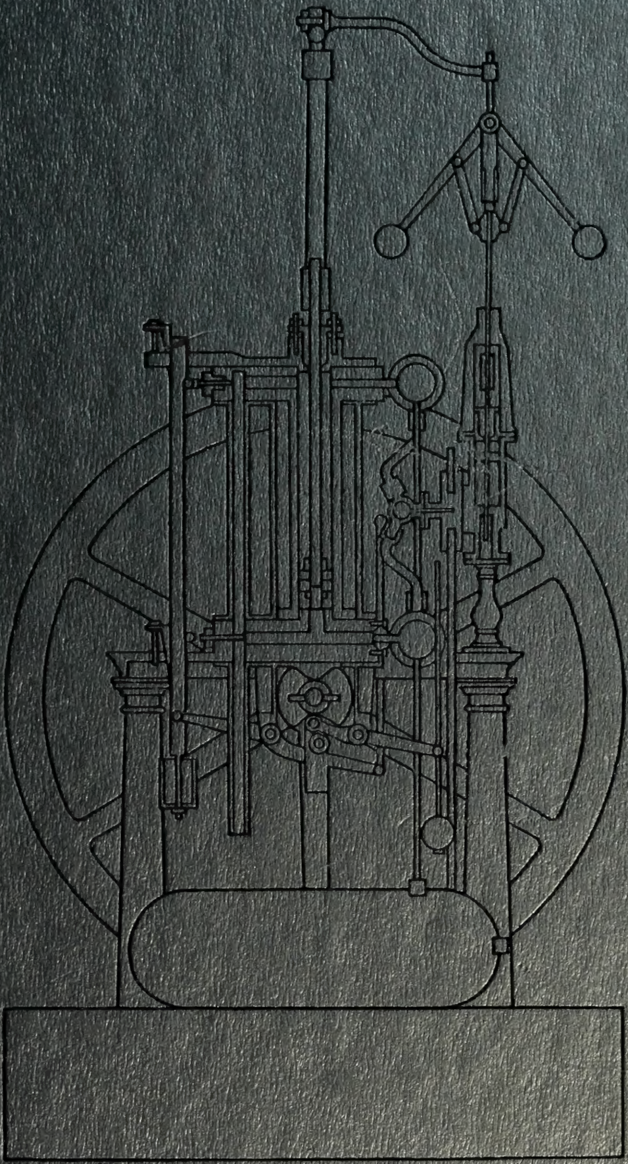


HOW THINGS WORK

HOW THINGS WORK

I

ILLUSTRATIONS RESEARCHED BY
ROGER JEAN SÉGALAT

Edito-Service S.A., Geneva, Publishers

INTRODUCTION

BY THE RIGHT HON. THE LORD RITCHIE-CALDER
Kalinga Prize winner for promoting the common understanding of science

"The world will never starve for want of wonders but only for the want of wonder..."—G. K. CHESTERTON

SO MUCH has happened since the end of World War II—the Atomic Age, the Computer Age and the Space Age, epochs concentrated into a brief lifetime—that wonder is liable to become numbed and wonders taken for granted. We flick a switch, press a button, or put a coin in the slot and expect things to happen...without asking "How?" That is sad, because as Sir Edward Appleton, the Nobel Prize-winning scientist, said, "It is such fun finding out."

And we cannot all be like my Arctic guide, Luke the Inquisitive Eskimo, who had a direct way of finding out. If he wanted to know how something worked—a wristwatch, for instance—he would just take it to pieces and memorize the mechanism. I found this childlike curiosity amusing until, when we were holed up in an igloo by a blizzard, I woke up one morning to find him dismantling my tape recorder.

Or one can do what I have done most of my life as a science writer: go to the chap who discovered something and ask him, "How does it work?" That is how, over the years, I learned about a lot of things in this book.

In going through the first volume of *The Way Things Work*, however, I was pleased (and sometimes professionally mortified) to find explanations of things which had slipped into the commonplace of my experience and about which I had never asked "How?" I was given one of the first practical ball-point pens in 1938 but forgot—all this time—to ask just exactly how it worked. It is simple, like most good ideas. For years I have been using a Polaroid camera. To photograph and develop immediately is like practicing a conjuring trick; it fascinates children and grownups alike; and sometimes it spoils the relish of a trick when one knows exactly how it is done. So I went on conjuring without asking! Now I know that the "How" is just as fascinating as the "Hey presto!"

Away back in 140 A.D., Hero of Alexandria invented the first coin-in-slot vending machine for dispensing holy water. It is interesting to have

it brought up to date in the modern mechanisms of the automats and the refined tortures of the jukebox. When they were first invented, I wrote about transistors, lasers, the magnetron, xerography, jet engines, hovercraft and computers, and it is nice to see them, grown-up, as it were, in the family album.

The bits and pieces of technology, however, are not enough. We would still be mystified, left with unanswered questions, if we did not see how they fitted into the patterns and the systems. If this and its preceding volume were just expanded glossaries of gadgets, they might suffice for a quiz game of "Ask me another." But this encyclopedia is designed to be more than that—it deals with principles and is grouped in subjects so that the picture of technology emerges and embraces the items.

I do not know who thought of *The Way Things Work*, nor how it was compiled, but I know it is a valuable piece of communication (to use an overworked word) to the general public and, speaking professionally, a useful crib for science writers and for scientists and technologists as well. It is simple without being condescending.

FOREWORD

THIS VOLUME is not a reference book in the ordinary sense. It has been designed, instead, to give the layman an understanding of *how things work*, from the simplest mechanical functions of modern life to the most basic scientific principles and complex industrial processes that affect our well-being. The result is, we believe, a unique book—a graphic and original introduction to the modern world of technology.

When it was originally published, in Germany, this book was called *Wie funktioniert das? (How Does It Work?)*. That question remains the key to the purpose and layout of the book. Here you will find the answer to the question posed by the inquisitive child who wants to know how a vacuum cleaner or a refrigerator works, or by you yourself, puzzled by the complexities of lasers or the secrets of Polaroid color photography. A page of descriptive text faces the coloured drawing to which it refers. Certain common principles have prompted the grouping on successive pages of related machines in which these principles are applied. Thus, for example, an entry on Light introduces a group of entries running in sequence from lenses and telescopes to cameras and colour television, etc. The various subjects are shown in the list of contents, overleaf, and may prove useful for cross-reference purposes. In addition, there is a full index at the end of the book, in which both machines and topics are listed alphabetically.

THE PUBLISHERS

CONTENTS

HOW THINGS WORK

When a liquid is heated, the molecules of the liquid begin to perform increasingly violent movements. They collide with other molecules which are not yet in violent motion, and in the end some of the molecules attain such high velocities that they burst through the surface film of the liquid (Fig. 1). They thus escape from the liquid as a gas. When they encounter the molecules of a colder medium, these escaped molecules part with some of their kinetic energy, coalesce, and finally form water droplets. This process can be observed when water is boiled in a kettle: the water vapour, or steam, which emerges from the kettle condenses against the kitchen windows. Similarly, water which is evaporated by the sun's heat will, on cooling when it rises to great altitudes, similarly condense to tiny droplets (clouds) which may eventually coalesce into larger drops that fall as rain. In chemical technology this process is performed in a distilling apparatus (Fig. 2) which, if it is of large size, is known as a distilling column (Fig. 3). In this way distillation is used for separating evaporable substances from non-evaporable ones, or separating substances which evaporate easily from others which evaporate less easily. Distillation is therefore a process for the separation of substances according to their volatility. The portion which evaporates when heat is applied and which is subsequently condensed by cooling it in coolers is called the distillate; the portion which does not evaporate is called the distillation residue. Distillation is one of the most important methods of purifying volatile substances and has been practised for thousands of years. High-efficiency distillation columns used in the chemical industry may be over 300 ft. high and as much as 16 ft. thick. They are employed more particularly in the petroleum industry for the production of power gases, fuels, lubricating oils, etc., but distillation is, for example, also an essential process in the manufacture of alcohol.

However, distillation can be used only for the separation of substances whose respective boiling points differ at least 0·5 – 1° C, unless additivies can be used to facilitate the process in special cases (extractive distillation). According to Raoult's Law, a proportion of the substance with the higher boiling point will—depending on its quantity and vapour pressure—be intermingled with the vapour of the substance with the lower boiling point, so that the latter substance will always to a greater or less extent be contaminated by the higher-boiling substance. By returning a portion of the distillate of the lower-boiling substance to the head of the column, however, the separation efficiency can be improved, as for example in rectifying installations (Fig. 3).

Every distilling plant comprises a source of heat and an evaporator (continuous-flow heater, still, distilling flask, retort). For the separation of more volatile mixtures a fractionating unit is attached to the evaporator, so that a distillation equilibrium is established depending upon the qualitative and quantitative composition of the mixture and the effect upon the temperature and pressure. Finally, the distillate is passed to a cooler, where it is cooled, and is then wholly or partially removed. In difficult separation processes a proportion of the distillate (depending upon the nature of the problem and the separation efficiency of the column) is returned to the column. Various forms of distillation are distinguished: one-way, vacuum, molecular, pressure and dry distillation. In the last-mentioned process a non-volatile substance, such as wood, is split up into distillable components by the application of heat.

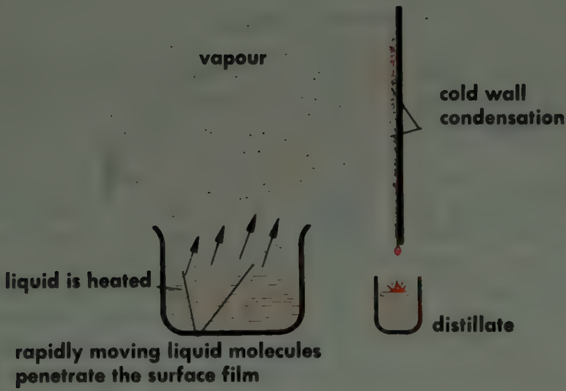


Fig. 1 PROCESSES OCCURRING IN DISTILLATION

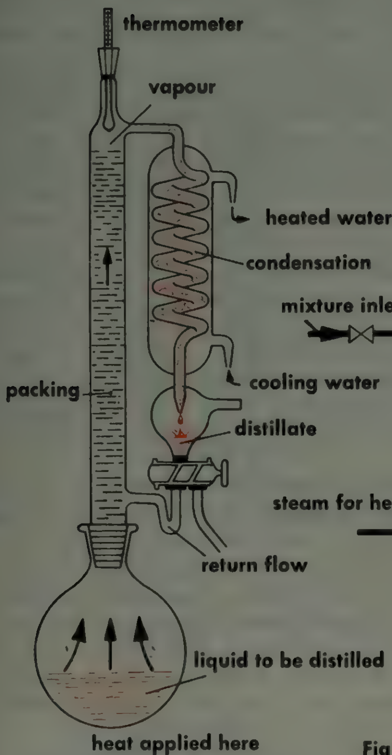


Fig. 2 MODERN DISTILLING APPARATUS

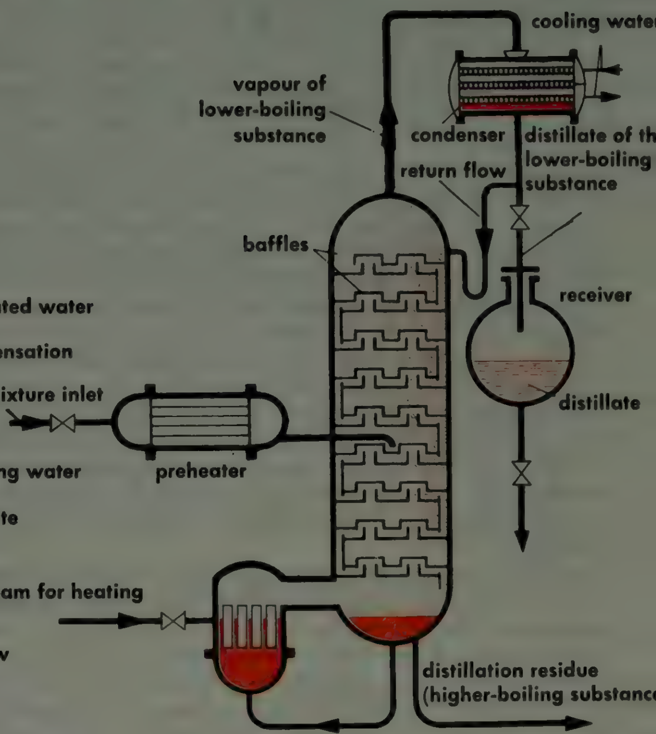


Fig. 3 CONTINUOUSLY-OPERATING DISTILLING COLUMN (RECTIFICATION)

A centrifuge is an apparatus intended mainly for the separation of heterogeneous mixtures (liquid-solid, e.g., sugar crystals in molasses) and liquid mixtures (e.g., liquids of different specific gravity which are not soluble in one another, such as fat in milk).

The physical principle on which the functioning of the centrifuge is based can be explained by first considering what happens when grains of sand suspended in water settle to the bottom. When sand is stirred up in a jar containing water, the force of gravity will exercise a more powerful attraction on each individual sand grain than on the water particles because a grain of sand has a greater mass than the surrounding water particles. The sand grains are therefore pulled down to the bottom of the jar. This is called sedimentation. After a time the system comes to rest; we then see that two layers have formed in the jar: on the bottom is the settled sand, and over this is the water. A similar separation can also be produced in a centrifuge. Indeed, it can be speeded up by subjecting the grains of sand (or other particles of matter suspended in the liquid) to a centrifugal force instead of gravitational force. Centrifugal forces come into play when a drum containing the mixture to be separated is rotated at high speed about its longitudinal axis. The magnitude of the centrifugal force, which is directed outward from the centre of rotation, is determined mainly by the speed of rotation of the drum—for the same reason that the iron ball which is being swung round and round (Fig. 2) will (for a given angle of elevation) fly farther, when it is released, in proportion as the athlete spins round faster.

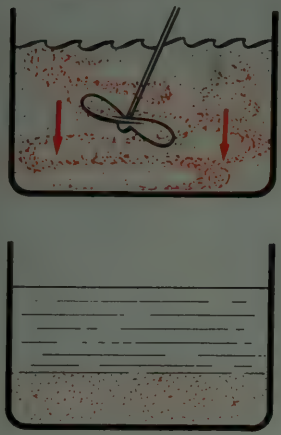
If the rotating drum is unperforated, the components of the mixture to be separated will, under the action of the centrifugal force, be deposited layerwise—according to the size of the particles and their specific gravity—against the wall of the drum. The materials with the highest specific gravity or whose particles have the greatest mass will, generally speaking, be sedimented closest to the wall of the rotating drum (Fig. 3). The separated layers can be removed from the latter through stationary tubes provided for the purpose.

Another type of centrifuge (intended more particularly for the continuous separation of the components of a liquid emulsion) is shown in Fig. 4a. The mixture is fed continuously through a spout into the rotating drum, which is equipped internally with conical plates arranged one above the other. The lighter liquid flows upwards along the backs of the plates and emerges from the centrifuge. The heavier liquid is flung to the periphery of the centrifuging drum, and the conically shaped casing of the drum deflects it upwards, where it emerges through another outlet (Fig. 4b).

In centrifuges operating on the filter principle the drum is perforated and may be covered with filter cloth. When the drum rotates, the solid particles are retained at the inner wall of the drum, whereas the liquid goes through the perforations and is thus hurled out of the drum and removed from the casing of the centrifuge (the well-known domestic spin dryer operates in this way). Fig. 5 shows an industrial centrifuge of this kind. Such a machine is usually run intermittently, i.e., it is filled with the mixture to be centrifuged; on completion of the treatment, the drum is stopped and the sedimented material is removed from the wall. In the continuous-action centrifuge for the separation of a solid from a liquid, the drum is provided internally with stationary scrapers which continuously remove the solid material.

As a rule, the drum is driven by an electric motor through a gear unit (worm gear, bevel gear) or belt which rotates the shaft on which the drum is mounted. At the point where this shaft enters the fixed casing of the machine a stuffing box (which forms a seal) must be provided.

Fig. 1 SEDIMENTATION OF A HEAVY SOLID IN A LIQUID



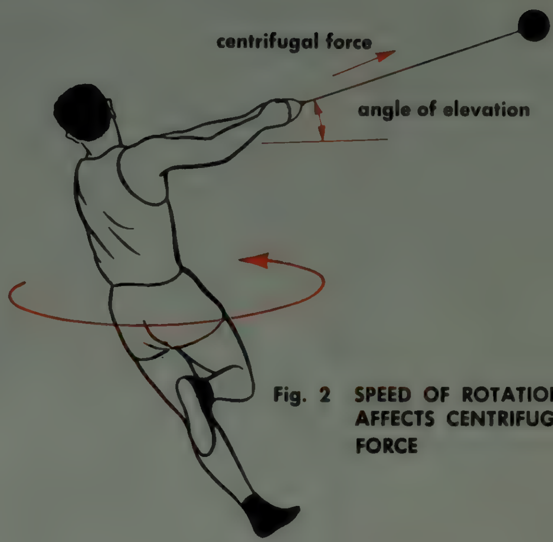


Fig. 2 SPEED OF ROTATION AFFECTS CENTRIFUGAL FORCE

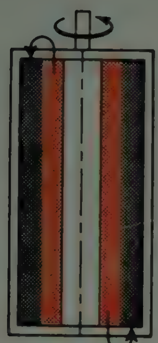


Fig. 3 SEDIMENTATION OF THE COMPONENTS OF A MIXTURE UNDER THE ACTION OF CENTRIFUGAL FORCE

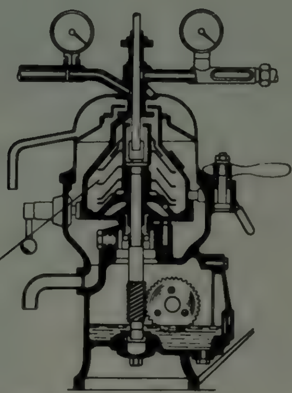


Fig. 4a CENTRIFUGE FOR THE SEPARATION OF TWO LIQUIDS

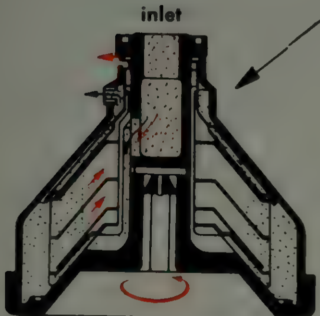


Fig 4b
CENTRIFUGING DRUM WITH PLATES
(*detail of Fig. 4a*)

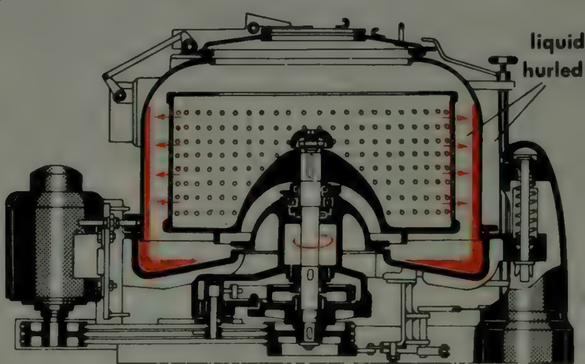
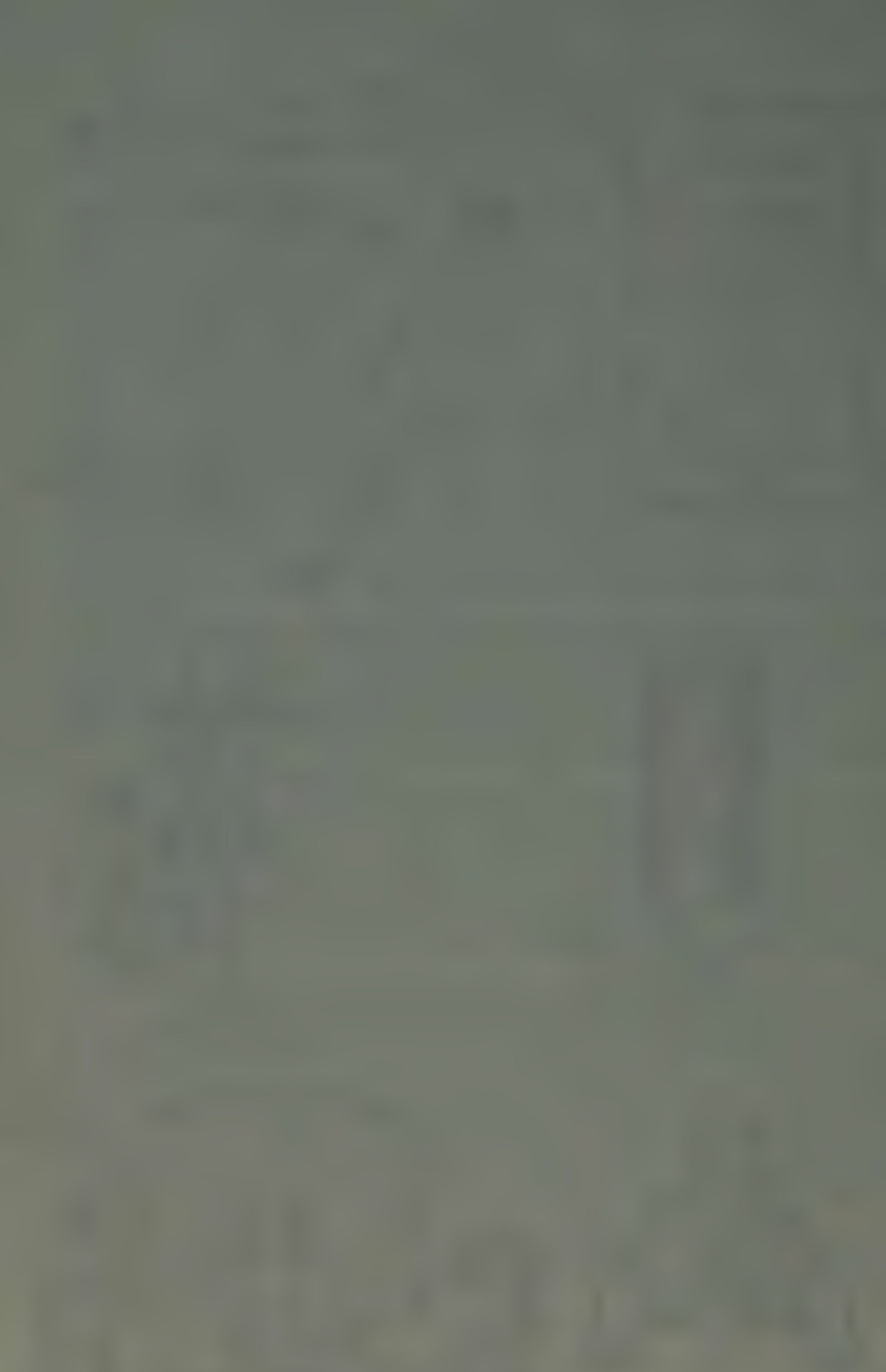


Fig. 5 FILTER DRUM CENTRIFUGE



Centrifuge used for the training of astronauts, Houston, Texas
Photo Cartier-Bresson, Magnum

Any ordinary combustion process is usually initiated by decomposition of the combustible material as a result of heating, which is usually confined to a small local area and which may be due to friction, irradiation, chemical oxidation, action of sparks or flames. This decomposition produces combustible gases which ignite because they react violently with the oxygen of the air. The reaction produces heat which causes further decomposition of the material and thus produces more gas. Eventually the temperature rises so high that the residue from the initial decomposition due to heat (namely, the carbon, which at first remains unaffected by the heat) also begins to burn, producing a large amount of heat in the process. This continuously liberated heat keeps the combustion process going until the combustible material, or the oxygen needed to sustain the process, has been consumed. Fire extinguishing agents must therefore either cool the combustible material, or they must cover this material with a firmly adhering non-inflammable coating, or they must rarify or displace the oxygen from the focus of the outbreak. As there are, for example, combustible substances which themselves contain the oxygen necessary to sustain combustion, and as it is of considerable importance—from the point of view of fire-fighting—whether the burning substance is liquid or solid, whether it is miscible with water, or whether it is lighter or heavier than the fire extinguishing agent, there must obviously be different kinds of extinguishing agents and different fire-fighting methods.

The carbon tetrachloride extinguisher (Fig. 1) contains anything from 0.5 to 6 litres[1] of this chemical (with certain additives), which is forced out of the extinguisher by the gas pressure from a cylinder of liquefied carbon dioxide. Carbon tetrachloride vaporises completely at 76.5° C, producing a heavy incombustible vapour. This kind of extinguisher is more particularly suitable for putting out fires in machinery and electrical installations. The carbon dioxide extinguisher (Fig. 2) contains 5 to 6 litres of carbon dioxide under high pressure. It is expanded to atmospheric pressure in the "snow tube", where the greater part rapidly vaporises and, in doing so, extracts so much heat from the surroundings that the rest of the carbon dioxide (about 30%) is cooled to solid carbon dioxide snow (—79° C). This "snow" is sprayed on to the fire by the carbon dioxide gas and causes a lowering of the temperature to below the ignition point; it also displaces oxygen. This is an all-purpose extinguisher.

The water type fire extinguisher (Fig. 3) contains 6 to 12 litres of water containing dissolved sodium bicarbonate. When the pin is struck, it shatters a flask of concentrated sulphuric acid inside the extinguisher. The acid reacts with the sodium bicarbonate, whereby a large quantity of carbon dioxide is evolved which forces the water at high pressure out of the discharge pipe. Because of its high heat of evaporation, water exerts a powerful cooling action; besides, the water vapour displaces oxygen. Water is useless and, indeed, dangerous as an extinguishing agent for fighting a fire in electrical installations or when inflammable solvents catch fire.

The dry chemical extinguisher (Fig. 4) contains 4 to 12 kg of sodium bicarbonate which is hurled into the blaze by the gas pressure developed by liquefied carbon dioxide or nitrogen. In the fire the sodium bicarbonate is decomposed into soda (which covers the combustible material with an air-excluding crust) and into water vapour and carbon dioxide (which displaces the oxygen and therefore smothers the fire). The application of this type of extinguisher will depend upon the nature and place of the fire (secondary damage). In the larger foam type fire extinguishers (Fig. 5) an air- or nitrogen-filled foam is produced. The foaming agent may, for example, consist of a decomposed protein substance (such as horn waste). The foam is stabilised by the admixture of urea, plastics, etc. It exercises a smothering and cooling action on the fire. Foam extinguishers can be used for any fire-fighting purpose, other than for putting out fires in machinery and electrical equipment.

1. One litre equivalent to 1.0567 liquid quarts.

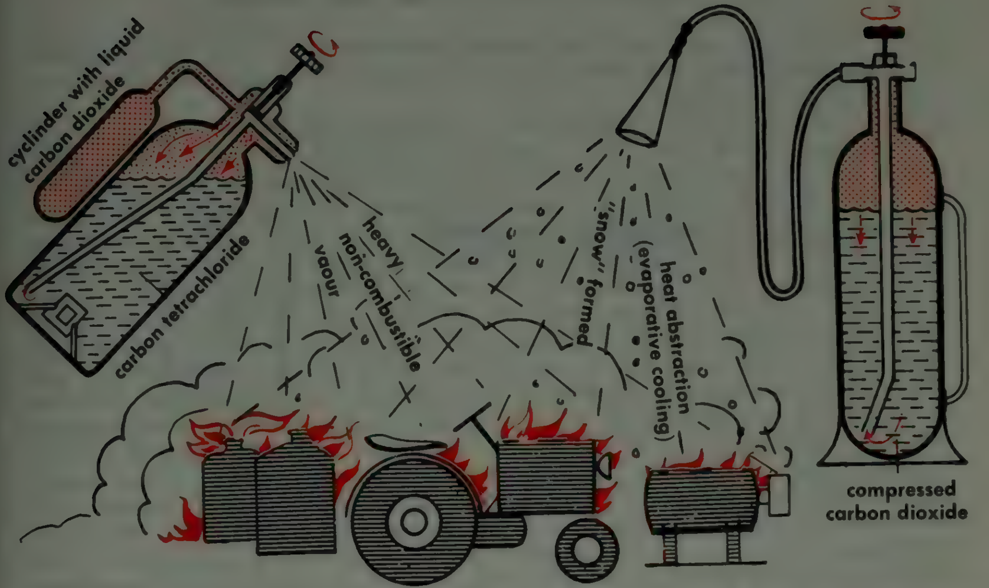


Fig. 1 CARBON TETRACHLORIDE EXTINGUISHER Fig. 2 CARBON DIOXIDE EXTINGUISHER

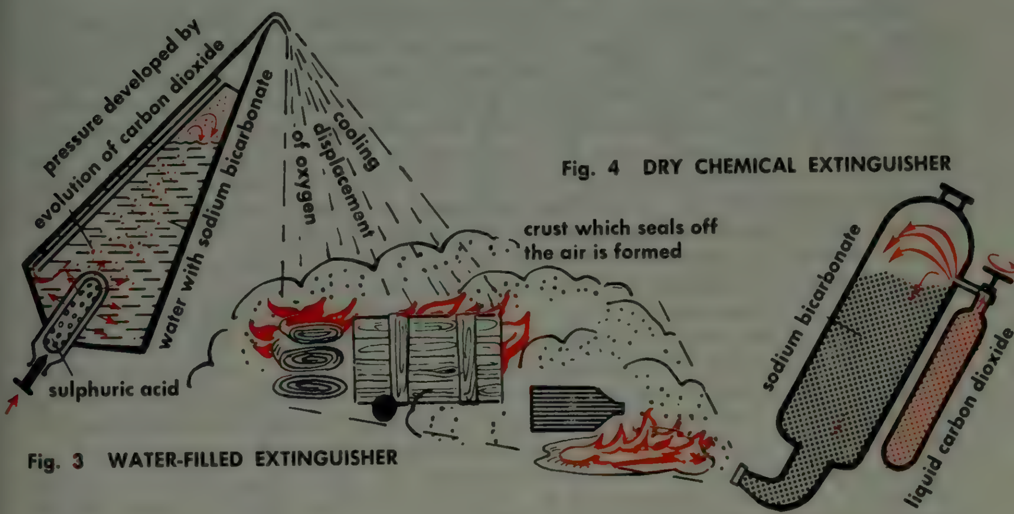


Fig. 4 DRY CHEMICAL EXTINGUISHER

Fig. 3 WATER-FILLED EXTINGUISHER

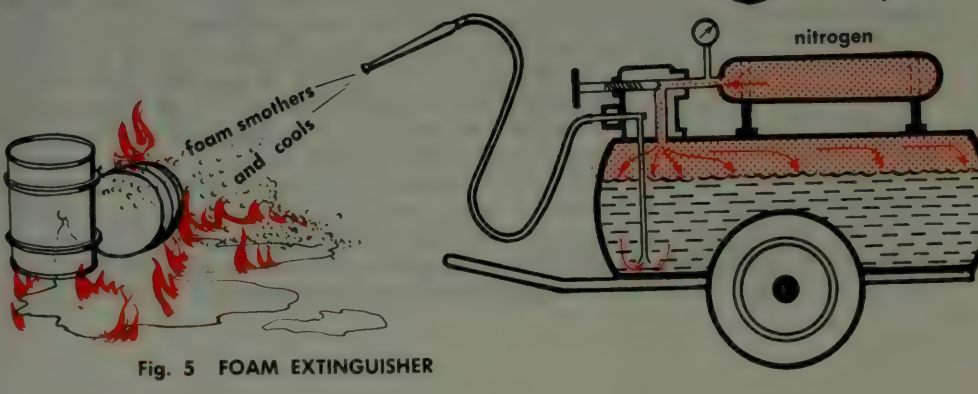


Fig. 5 FOAM EXTINGUISHER

The functioning principle of an ordinary *thermometer* is based on the property of thermal expansion possessed by most substances, i.e., they expand when heated and contract on cooling.

The commonest type of thermometer is the mercury thermometer. It consists of a capillary tube (a tube with a very small bore) which is sealed at its upper end and is enlarged into a spherical or cylindrical bulb at its lower end. This bulb is filled with mercury. When this is heated, it expands and rises in the tube. Because of the very small bore of the latter, even a small increase in the volume of the mercury will cause it to rise quite appreciably. The thermometer is calibrated between two fixed reference points: the freezing point and the boiling point of water at normal atmospheric pressure (760 mm mercury; see page 252). The difference in level in the mercury column between these two points is divided into 100 equal parts, each division being one degree on the centigrade scale (1° C). On the Fahrenheit scale the difference between the two fixed reference points is divided into 180 parts, each division being one degree Fahrenheit (1° F), the freezing point and boiling point of water on this scale being 32° F and 212° F respectively. On the Réamur scale the corresponding temperatures are 0° R and 80° C, i.e., this scale is divided into 80 degrees between freezing and boiling. Mercury is not suitable for thermometers used for measuring very low temperatures, as mercury itself solidifies at −39° C. Such thermometers are filled with coloured alcohol, which has a substantially lower freezing point. The lowest conceivable temperature is the absolute zero, corresponding to −273.16° C. Absolute temperature is measured with respect to the absolute zero in degrees Kelvin, i.e., −273.16° = 0° K.

Thermocouple: If the ends of two wires of dissimilar metals or metal alloys (e.g., copper and constantan or copper and iron) are soldered together (Fig. 2a) and if one soldered junction is kept at a constant temperature, while the other junction is heated, a thermo-electric potential difference develops between the two junctions. This potential difference (voltage) is greater according as the difference in temperature between the junctions is greater and can be read on a voltmeter (Fig. 2b), which can be calibrated to give temperature readings. An arrangement of this kind is called a thermocouple. The voltage produced by a single thermocouple is very small (a few millivolts). To obtain a higher voltage, a number of thermocouples can be arranged in series (Fig. 3), with alternate hot and cold junctions. In this way a so-called thermopile, or thermo-electric battery, is obtained.

The *resistance thermometer* (Fig. 5) is a device whose operation depends upon the variation of the electrical resistance of a wire with temperature. Most metals become more resistant to the passage of an electric current as they become hotter, the increase in resistance being (within certain limits) proportional to the rise in temperature. The resistance used in the thermometer consists of a platinum or nickel wire and is so designed that its resistance at 0° C is 100 ohms. The variations in resistance due to temperature changes are measured as variations in the strength of a current by means of, for example, a crossed-coil instrument, whose pointer deflects by an amount governed by the ratio of the currents flowing through the two coils. The current in one of the coils is kept constant by means of a resistance which is unaffected by temperature; the current in the other coil is determined by the resistance of the thermometer wire which varies with the temperature.

A *bimetallic thermometer* comprises two strips of dissimilar metals soldered together. These metals have different coefficients of thermal expansion and therefore undergo different increases in length on heating. Fig. 4 illustrates the functioning principle of a temperature measuring device embodying a bimetallic spiral whose curvature varies with the temperature and causes a pointer to deflect. The scale is calibrated by establishing the positions of the pointer at certain known temperatures and then marking the scale so that each division corresponds, say, to one degree.

Fig. 1
THERMOMETER

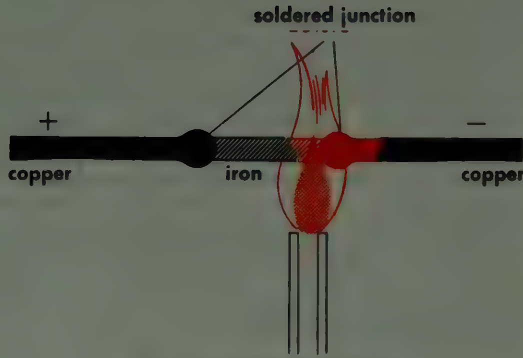


Fig. 2a THERMOCOUPLE

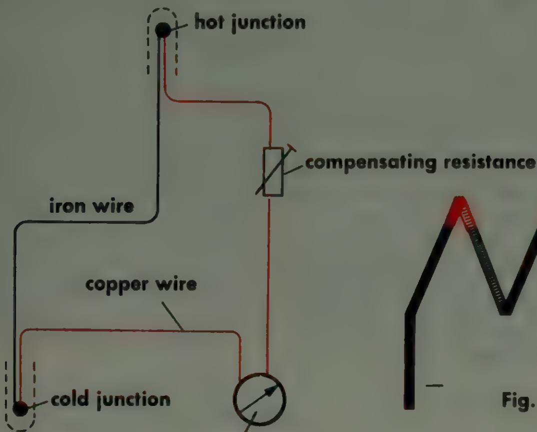


Fig. 2b
ELECTRICAL CIRCUIT FOR TEMPERATURE MEASUREMENT WITH THERMOCOUPLES

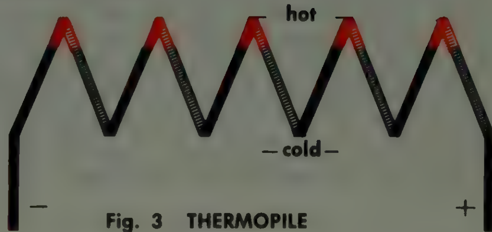


Fig. 3 THERMOPILE

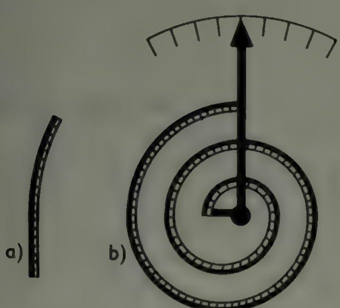


Fig. 4a BIMETALLIC STRIP CURVED BY CHANGE IN TEMPERATURE

Fig. 4b BIMETALLIC STRIP AS THERMOMETER

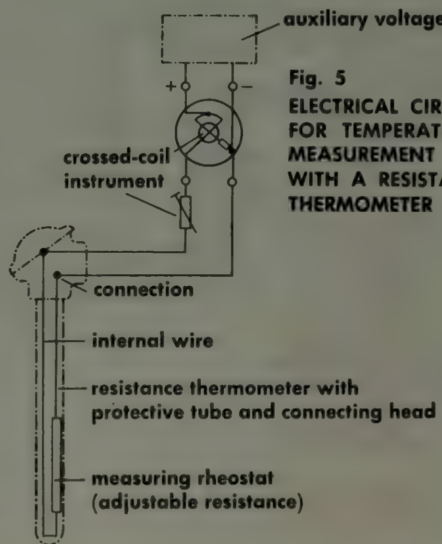


Fig. 5
ELECTRICAL CIRCUIT FOR TEMPERATURE MEASUREMENT WITH A RESISTANCE THERMOMETER

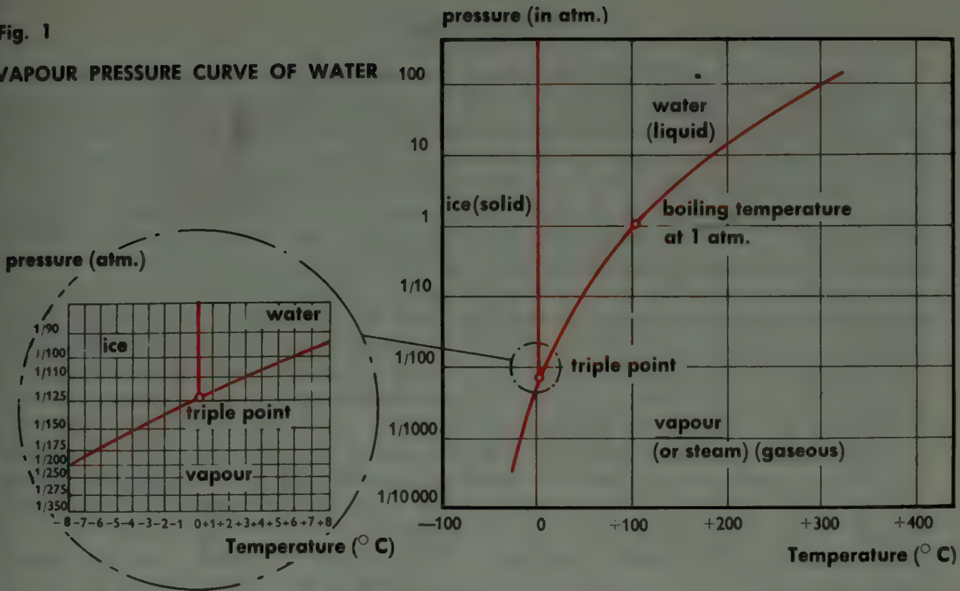
"Dry ice" is a name sometimes applied to compressed carbon dioxide "snow", i.e., solid carbon dioxide with a temperature of $-79°$ C. Under normal conditions carbon dioxide is a colourless and odourless gas with a density about $1\frac{1}{2}$ times as high as that of air. Like water (and indeed all other substances), it can occur in the gaseous, liquid or solid state, depending on the physical conditions. In addition, carbon dioxide possesses the property of sublimation, i.e., it can pass directly from the solid to the gaseous state without becoming liquid.

Under which conditions of pressure and temperature a particular state occurs is indicated by the vapour pressure curve (Figs. 1 and 2). In accordance with the lines in the vapour pressure diagram, the transition from one state to the other, in conjunction with absorption or release of heat. Of particular interest is the so-called triple point, where all the three states—gaseous, liquid, solid—can co-exist simultaneously. For example, the triple point for water is located at a pressure of 4.6 mm mercury column (approx. 6/1000 atm.; Fig. 1) and 0.01° C. On the other hand, for CO_2 this point is located at 5.1 atm. and $-56.2°$ C. The vapour pressure curve indicates the relationship between the boiling point of a substance and the pressure. Thus, water (Fig. 1) boils at 100° C at a pressure of 1 atm. (= normal atmospheric pressure); carbon dioxide (Fig. 2) "boils" at 0° C at a pressure of 60 atm.

To make carbon dioxide snow, carbon dioxide is cooled at high pressure (up to 70 atm.) and liquefies in consequence. Further cooling takes the carbon dioxide to the triple point. Now the compressed liquid carbon dioxide is suddenly expanded by spraying and turns into "snow". This happens because the evaporation of part of the liquid causes intensive cooling of the rest (see page 46). The dividing line between liquid and solid in Fig. 2 is crossed: the carbon dioxide turns from liquid to solid. To achieve this result, the carbon dioxide gas is liquefied by means of three- or four-stage compressors (see page 56) with intermediate and final cooling, the liquid carbon dioxide then being expanded in a tower (on the right in Fig. 3). About one-third of the liquid is thereby transformed into snow; the other two-thirds turn into gas, which is removed by suction, recompressed, and returned to the process. The snow is pressed into blocks weighing 50–250 lb.

Fig. 1

VAPOUR PRESSURE CURVE OF WATER



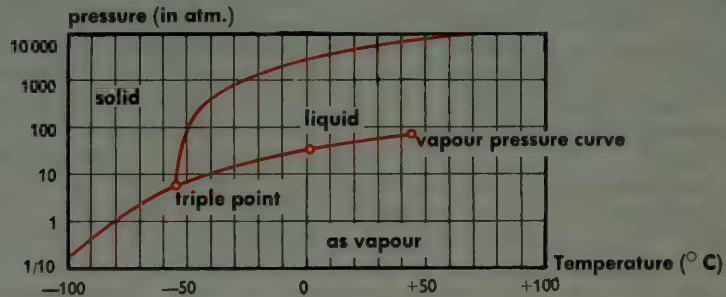


Fig. 2 VAPOUR PRESSURE CURVE OF CARBON DIOXIDE (CO_2)

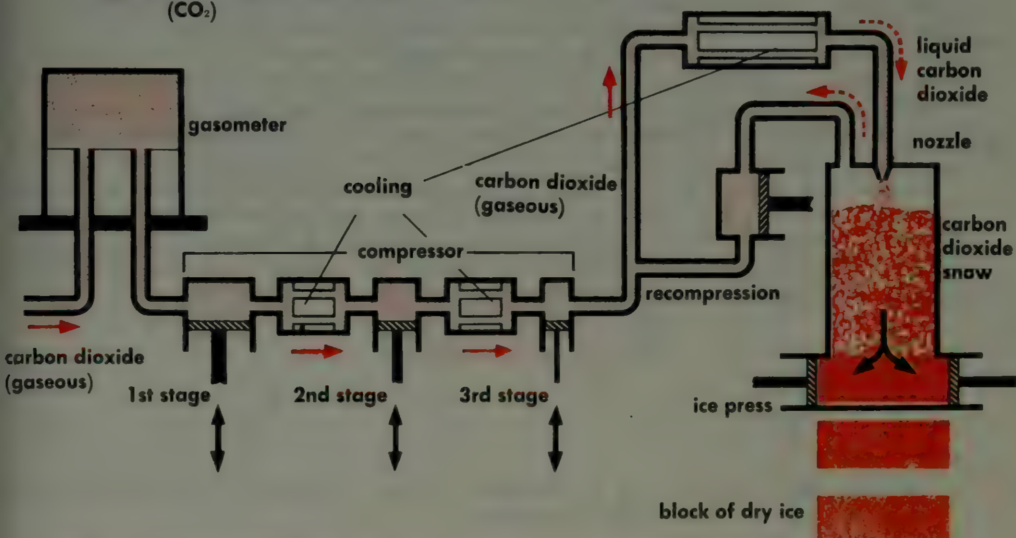


Fig. 3 MANUFACTURE OF DRY ICE
(*schematic*)

A thermostat is a device for maintaining a temperature constant at a desired value. For this purpose it is equipped with a temperature sensing unit which detects any deviation of the actual temperature from the desired value and transmits information on this to a device which cancels the deviation. The sensing unit may be a tube filled with a liquid, a bimetallic strip or a spring bellows. The simplest device of this kind is the *direct*-acting thermostat. It makes use of the fact that nearly all liquids expand on heating (Fig. 1). The thermostat itself consists of a tube filled with a liquid which expands very considerably when it is heated (Fig. 2). The connection to the control device which actuates the valve in the hot-water supply pipe (for example) is established by a capillary tube which is also filled with liquid. If the air temperature in the room under thermostatic control rises above the desired level, the liquid in the sensing unit expands, overcomes the restraining force of a spring on the valve, and throttles or closes the latter. As a result of this, the flow of hot water (or other heating medium) is reduced, and less heat is supplied to the room. Because of this the temperature in the room will go down after a time, so that the liquid cools and contracts. The spring load on the valve once again exceeds the pressure exerted by the liquid and opens the valve. In this way the temperature in the room is kept constant within fairly narrow limits. The desired value of the temperature is set on a graduated scale which has been calibrated by the makers. By rotation of the screw on the control device the valve spring is compressed to a greater or less extent by the liquid, so that the valve correspondingly opens more or less. The flow rate of the hot water thus increases or decreases, causing the temperature level in the room to rise or to fall (Fig. 2). A different type of device is the *indirect*-acting thermostat, which uses an auxiliary source of power (e.g., electricity or compressed air) for transmitting the impulses for effecting the change in temperature. In a device of this kind the sensing unit actuates a contact, i.e., an electrical contact is closed (Fig. 3), so that, for example, an electromagnetically controlled valve of the heating system is opened or closed. For instance, if the above mentioned liquid-filled sensing unit is employed, the expansion of the liquid due to a rise in temperature closes the contact. The electric current which then energises the electromagnetic valve will close the latter. When the temperature falls, the reverse process takes place. When the pre-set minimum temperature is reached, and the liquid in the sensing unit has contracted a certain amount, the circuit is broken, so that the electromagnet now ceases to keep the valve closed. Consequently the hot water flows through the system once again.

Another type of sensing unit is the so-called bimetallic strip (Fig. 4b). This consists of two thin strips of different metals bonded one against the other to form a composite strip. These metals undergo different amounts of thermal expansion. One end of the bimetallic strip is fixed. When the temperature rises, one metal expands more than the other, causing the strip to curve, so that its free end actuates a contact (Fig. 4a).

A third type of sensing unit consists of a resilient bellows filled with a volatile liquid or a gas (Fig. 4c). When the temperature rises, the increase in volume of the vapour or gas in the bellows causes the latter to expand and actuate a contact. A fall in temperature has the opposite effect, i.e., the bellows contracts and may (as in the example illustrated) actuate another contact.

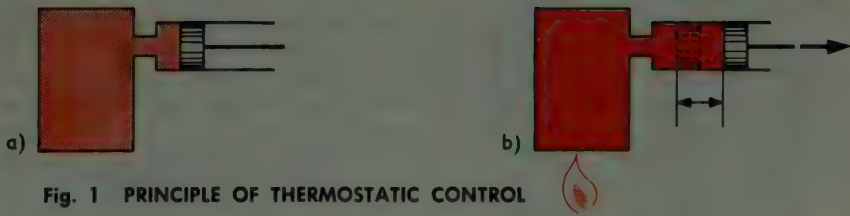


Fig. 1 PRINCIPLE OF THERMOSTATIC CONTROL

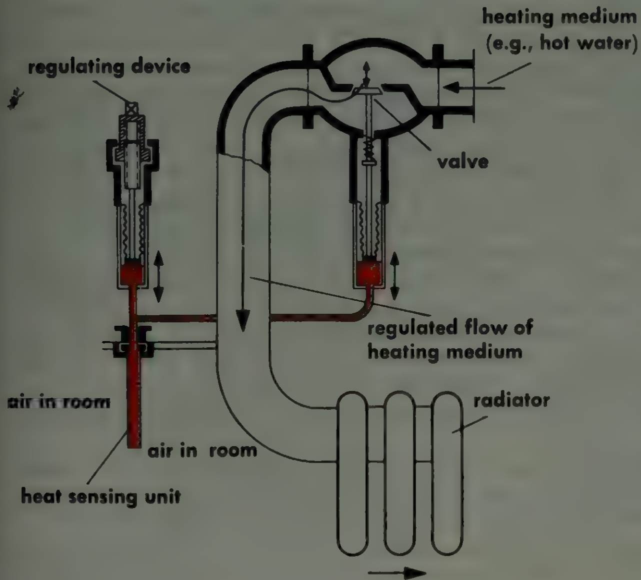


Fig. 2 THERMOSTAT FOR SPACE HEATING CONTROL
(*direct action*)

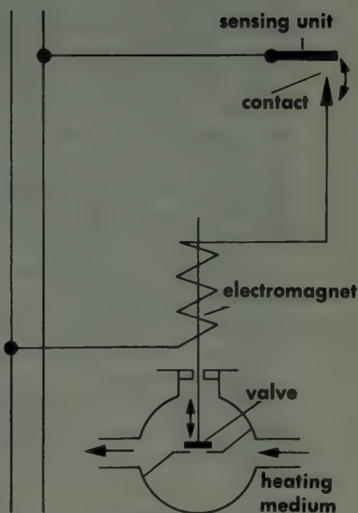


Fig. 3 DIAGRAM OF A
DIRECT-ACTING THERMOSTAT

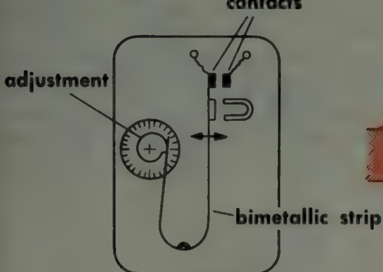


Fig. 4a BIMETALLIC SENSING UNIT

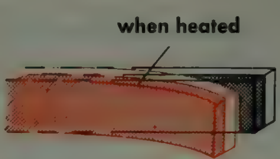


Fig. 4b PRINCIPLE OF THE
BIMETALLIC STRIP

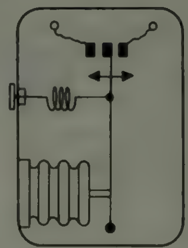


Fig. 4c BELLOWS-TYPE
SENSING UNIT

Presses provide a means of compressing and shaping components by exerting a high pressure on them. Depending on their method of functioning, a distinction is to be made between hydraulic and mechanical presses. *Hydraulic presses* use a gas or a liquid (usually water) as the power-transmitting medium. Their operation is based on the phenomenon that the pressure exerted on a liquid or gaseous medium which is compressed in a cylinder is of the same magnitude at all points of the cylinder. "Pressure" is defined as force acting per unit of area. If the liquid is compressed by means of a small piston in a small cylinder which is in communication with a large cylinder closed by a large piston (as in Fig. 1), then the situation will be as follows: on one side a small force acts upon a small area (i.e., the small piston) and produces a certain pressure. On the other side this pressure acts upon a large area (i.e., the large piston). A large force can therefore be developed by the large piston (since pressure is force per unit of area, then force must be pressure times area). The volume that is displaced by these two piston movements is, however, the same in both cylinders, i.e., the small piston has to travel a great distance to make the large piston move only a short distance. A hydraulic press is operated, not by one, but usually by three small pistons which consecutively force water into the large cylinder. To ensure that the pulsation of the liquid flow is not transmitted to the main cylinder of the press, the delivery pipe is provided with an air vessel which is partly filled with air and has a cushioning effect (Fig. 2). Normally this air has the same pressure as that in the delivery pipe. If a sudden surge of pressure occurs, the air is further compressed and absorbs the rise in pressure. Conversely, if the quantity of hydraulic medium delivered through the pipe is inadequate to meet brief high demands, the air pressure in the air vessel can to some extent compensate for this by boosting the pressure of the liquid in the pipe.

When the water pump (i.e., the small pressure generating cylinder) is started up, it fills the main cylinder with water and produces a pressure which causes the ram of the press to descend and exert pressure on the work-piece placed under it. Retraction of the ram is effected by two small pistons actuated by water under relatively low pressure.

Mechanical presses have various drive systems. In the screw press (Fig. 3a) a screw spindle is rotated in a fixed nut, whereby a longitudinal force is developed in the spindle, one end of which thrusts against the work-piece laid under it.

The upper end of the spindle of a simple manually-operated press as illustrated in Fig. 3b (known as a "fly press") is provided with a cross-piece for rotating it.

On the larger presses the upper end of the screw spindle is provided with a large flywheel which, when it is rotating, contains a large reserve of stored-up energy. This energy is transmitted through the spindle to the work-piece. The flywheel is driven by a friction wheel. Retraction is effected by another friction wheel, which drives the flywheel in the opposite direction (Fig. 3a).

A further modification of the mechanically driven press is the eccentric press (Fig. 4) and the crank press (Fig. 5). In both cases a large flywheel is driven by a motor or other prime mover. When the ram of the press strikes the work-piece, the rotational energy of the flywheel produces a torque (twisting moment) in the shaft, whereby a relatively large force is developed at the eccentric or at the crank. This force is imparted to the work-piece.

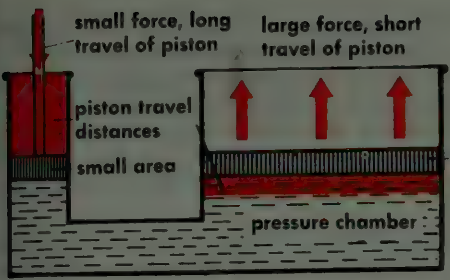


Fig. 1 PRINCIPLE OF THE HYDRAULIC PRESS

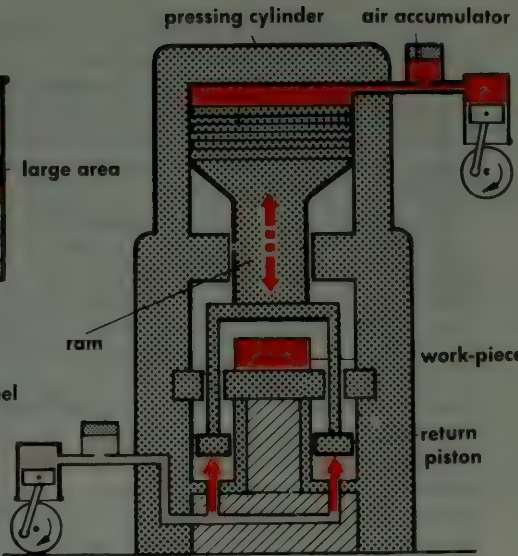


Fig. 2 HYDRAULIC PRESS (*schematic*)

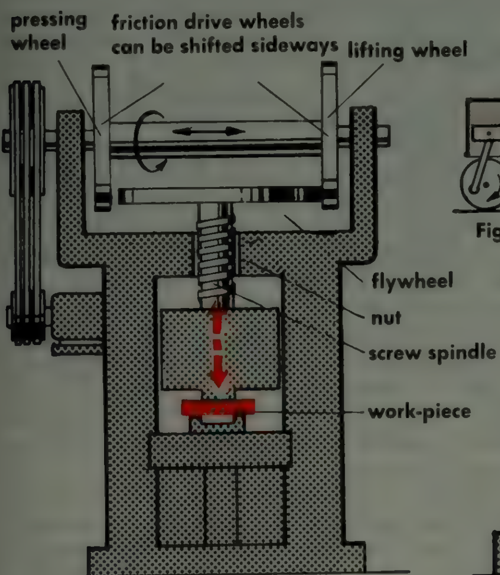


Fig. 3a SCREW PRESS (*schematic*)

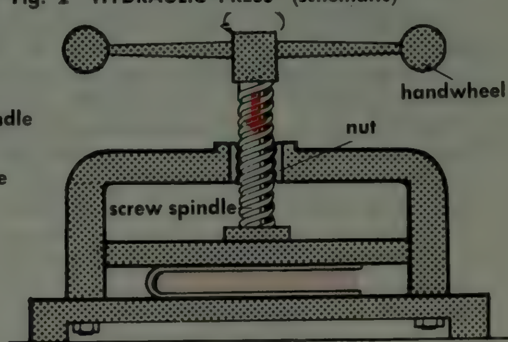


Fig. 3b SCREW PRESS WITH HANDWHEEL (FLY PRESS)

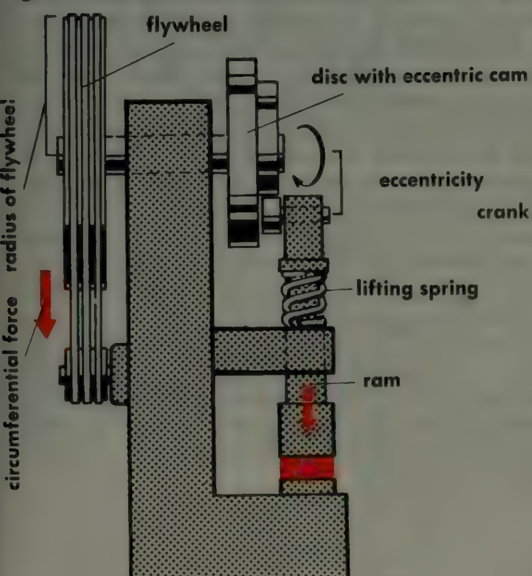


Fig. 4 ECCENTRIC PRESS (*schematic*)

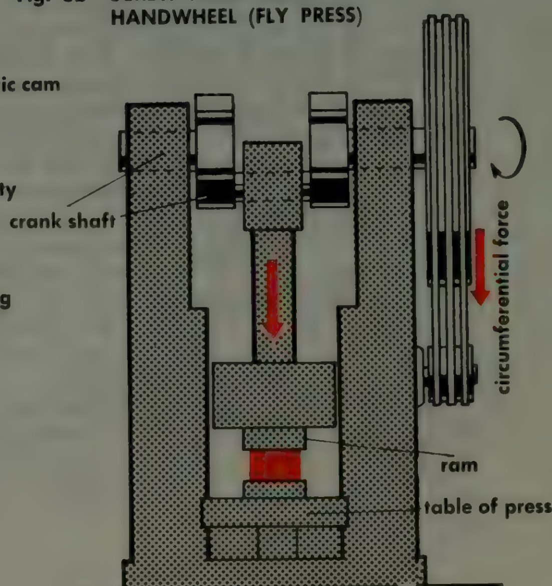


Fig. 5 CRANK PRESS (*schematic*)

Pumps are used for the transport of liquids or gases through pipes. The general principle is as follows: on one hand, a suction is produced, whereby the liquid is drawn in; on the other hand, an excess pressure to overcome the counter pressure is developed, causing the fluid (liquid or gas) to be forced away. According to the various operating principles, several types of pump are to be distinguished: 1. Piston pumps, which operate with reciprocating (up-and-down or to-and-fro) or rotating pistons. 2. Centrifugal pumps, which operate with rotating blades. 3. Jet pumps, which utilise the energy of flowing fluids.

1. *Piston pumps*: Consider a piston in a cylinder, as in Fig. 1. When the piston moves to the right from its extreme left initial position, water flows into the vacant space that is formed in the cylinder. This happens because the atmospheric pressure acting upon the surface of the water outside the pump forces water up the pipe into the cylinder, in which the outward stroke of the piston has produced a vacuum (or, at any rate, a lowering of pressure). Theoretically the atmospheric pressure can sustain a column of water 33 ft. in height. A pump would thus have a "suction head" of 33 ft., i.e., it would be able to draw water up to a height of 33 ft. into the cylinder. However, the fact that water has a certain vapour pressure and encounters resistance in the suction pipe, reduces the actual suction head to usually somewhere around 23 ft. (Fig. 2). When the piston has travelled to its farthest position, it changes its direction and starts to move back to the left. This produces a higher pressure in the cylinder and causes the inlet valve to close and the delivery valve to open. The latter valve remains closed during the suction stroke of the piston, as the pressure in the delivery pipe is higher than in the cylinder. The piston now pushes the liquid out of the cylinder, against the pressure in the delivery pipe. Special forms of the piston pump are the lifting pump, which discharges the liquid when the piston rises (Fig. 3); the diaphragm pump, in which a flexible diaphragm takes the place of the piston, this diaphragm being moved to and fro by an actuating rod (Fig. 4); and the vane pump, in which instead of a piston there is a vane which is swung to and fro in a circular casing (Fig. 5); the inlet valves are installed in a fixed partition of the casing, while the delivery valves are in the vane.

2. *Centrifugal pumps*: A centrifugal pump is very similar to a Francis turbine (see page 62) operating "in reverse". In a turbine the water drives a runner wheel[1], whereas in a pump a similar wheel fitted with vanes and known as an impeller imparts motion to the water (or other fluid). A piston pump delivers the water in a pulsating stream, but a centrifugal pump gives a steady flow. The pressure for achieving the required "delivery head" is produced by centrifugal acceleration of the fluid in the rotating impeller. The fluid flows axially towards the impeller, is deflected by it, and flows out through the aperatures between the vanes. Thus the fluid undergoes a change in direction and is accelerated. This produces an increase in pressure at the pump outlet. On leaving the impeller, the fluid may first pass through a ring of fixed vanes which surrounds the impeller and is known as a diffuser. In this device with its gradually widening passages the velocity of the liquid is reduced, its kinetic energy being converted into pressure energy. This conversion is completed in the volute of the pump, i.e., the gradually widening spiral casing (Fig. 6). It should be noted that in some pumps there is no diffuser, the fluid passing directly from the impeller to the volute. Pumps of the kind described here cannot, however, produce high delivery heads. A higher delivery head can be obtained by means of a multistage centrifugal pump in which two or more impellers are mounted one behind the other (compare Fig. 8 with Fig. 7).

As a rule, centrifugal pumps are not self-priming, i.e., they are unable to draw in the fluid on first being started up, because when the impeller is revolving in air in the "empty" casing, it cannot develop sufficient suction. These pumps therefore have to be primed (filled with water) for starting.

1. Rotor in U.S.A.

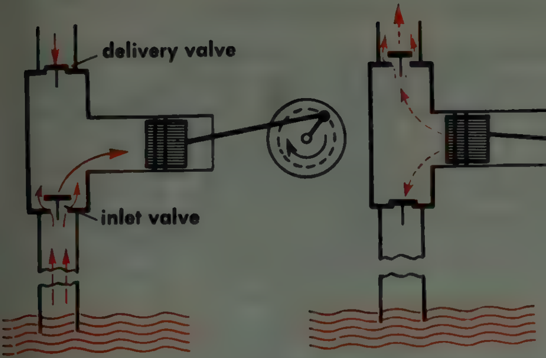


Fig. 1 OPERATING PRINCIPLE OF A PISTON PUMP

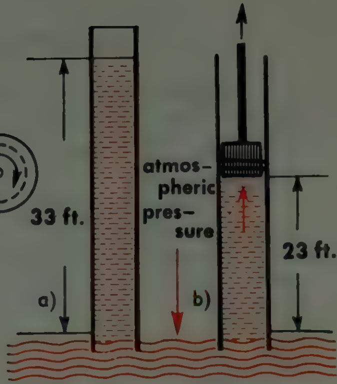


Fig. 2 THEORETICAL (a) AND ACTUAL (b) SUCTION HEAD

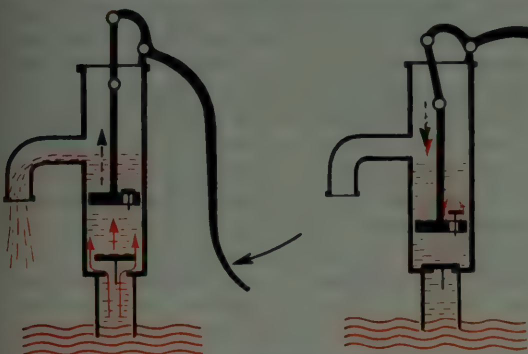


Fig. 3 LIFTING PUMP

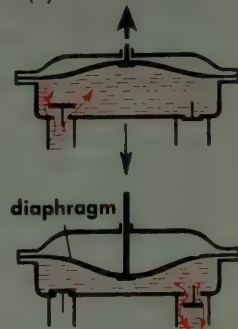


Fig. 4. DIAPHRAGM PUMP

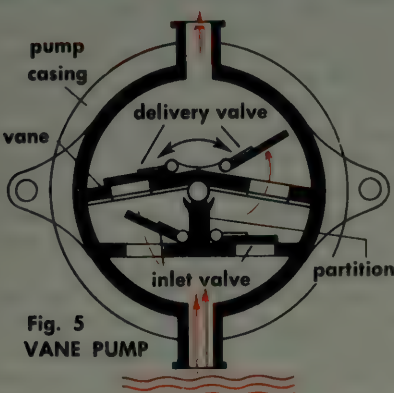


Fig. 5
VANE PUMP

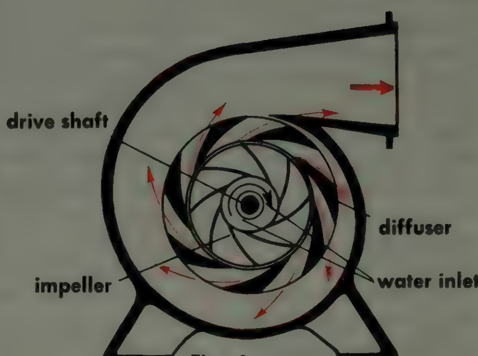


Fig. 6
CENTRIFUGAL PUMP
(*schematic*)

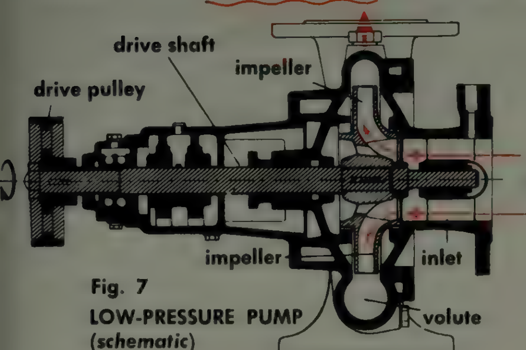


Fig. 7
LOW-PRESSURE PUMP
(*schematic*)

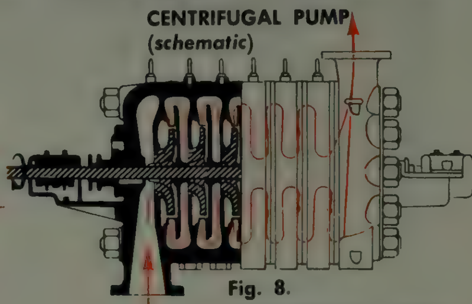


Fig. 8.
HIGH-PRESSURE PUMP
(*schematic*)

Compressors are machines for the compression of gases and vapours (to pressures of 2000 atm. and higher). Piston compressors are used for producing the highest pressures, whereas centrifugal compressors (see page 298) are used for low to medium pressures.

Fig. 1 shows a piston compressor. The crankshaft is driven by a suitable prime mover (electric motor, steam engine, internal combustion engine, etc.). A connecting rod between the crankshaft and the piston transforms the rotary motion of the former into a to-and-fro motion of the piston in the cylinder. The valves are spring-loaded and react to variations in pressure produced by the piston movement. In performing the suction stroke the piston causes a lowering of pressure in the cylinder, so that the inlet valve (Fig. 2) opens against the restraining pressure of its spring and allows the gas to flow into the cylinder. Then, when the piston begins to form its return stroke (compression stroke), the inlet closes because of the increase in pressure within the cylinder. When the piston has completed this stroke, i.e., has arrived at "top dead centre", the pressure of the gas compressed in the cylinder is so high that the delivery valve opens against the restraining pressure of its spring, allow the gas to flow into the delivery pipe. The spring of the delivery valve need not be as powerful as might at first be supposed. Actually there is also a pressure acting in the delivery pipe, at the rear of the valve, and this pressure is only a little below that which is produced by the compression stroke of the piston. Also, it is this pressure in the delivery pipe that keeps the delivery valve closed during the suction stroke. Usually, the compressor feeds the compressed gas into an intermediate vessel—called a receiver in the case of an air compressor—from which it is supplied to consumer equipment.

The operating cycle of a compressor is characterised by the so-called indicator diagram (Fig. 3). It shows the relation between pressure and volume during one revolution of the crankshaft, i.e., during one to-and-fro movement of the piston. Line 1 represents the suction stroke: the gas volume in the cylinder increases while the (low) pressure remains constant. Line 2 represents the compression: the pressure rises while the volume decreases. Then comes the discharge of the compressed gas, represented by line 3: the volume of the cylinder decreases at constant (high) pressure. Line 4 shows that when the piston performs the suction stroke (i.e., travels to the right), the inlet valve cannot immediately open, as the residual gas in the cylinder at the end of the compression stroke must expand until the pressure in the cylinder has fallen below the pressure in the inlet pipe. Not till then does the inlet valve open. Since the piston must not actually strike the end of the cylinder, there is always some dead space and therefore an undesirable gas residue, so that only a portion of the suction stroke is really utilised for drawing gas into the cylinder. If high pressures are required, compression will have to be achieved in several stages, i.e., the gas is passed from one cylinder to the next (Fig. 4). As gases become heated when they are compressed, it is necessary to provide cooling between the successive stages (Fig. 4a). These multi-stage compressors are usually of the double-acting type, i.e., when the piston travels in one direction, the gas is compressed on one side and suction is produced on the other side of the piston: the opposite occurs during the return stroke.

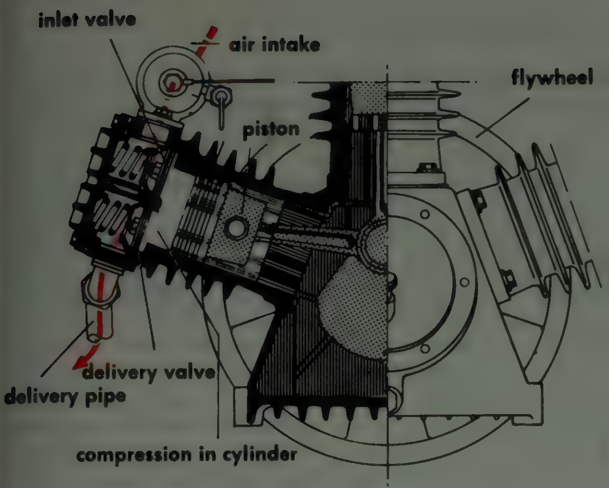


Fig. 1 SECTION THROUGH A COMPRESSOR

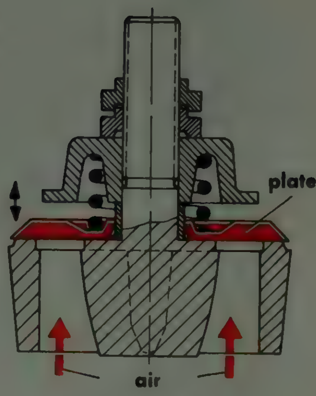


Fig. 2 VALVE

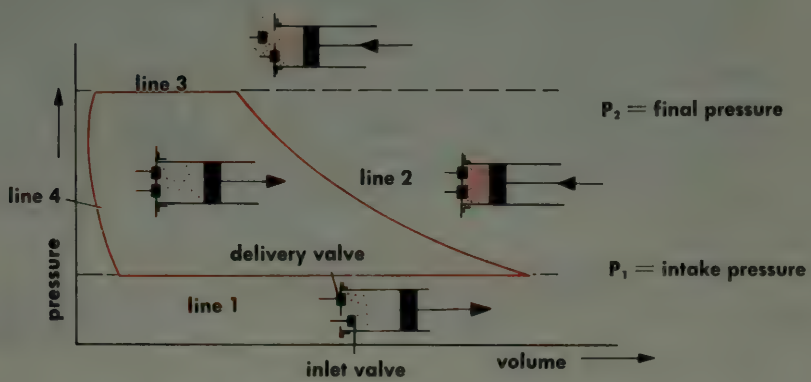


Fig. 3

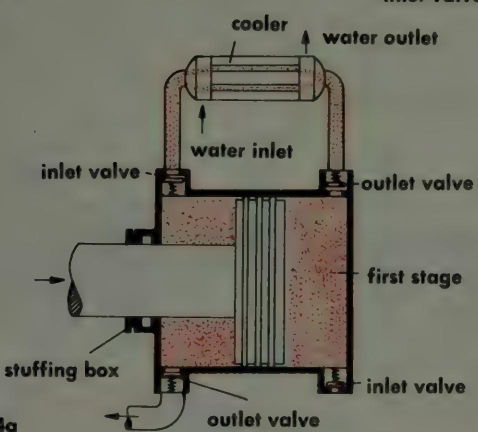


Fig. 4a

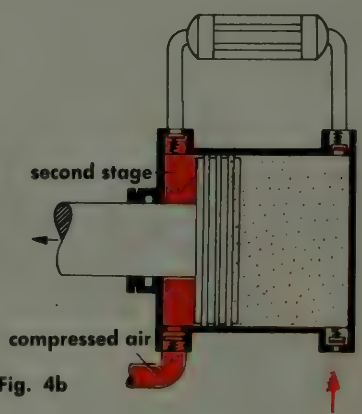


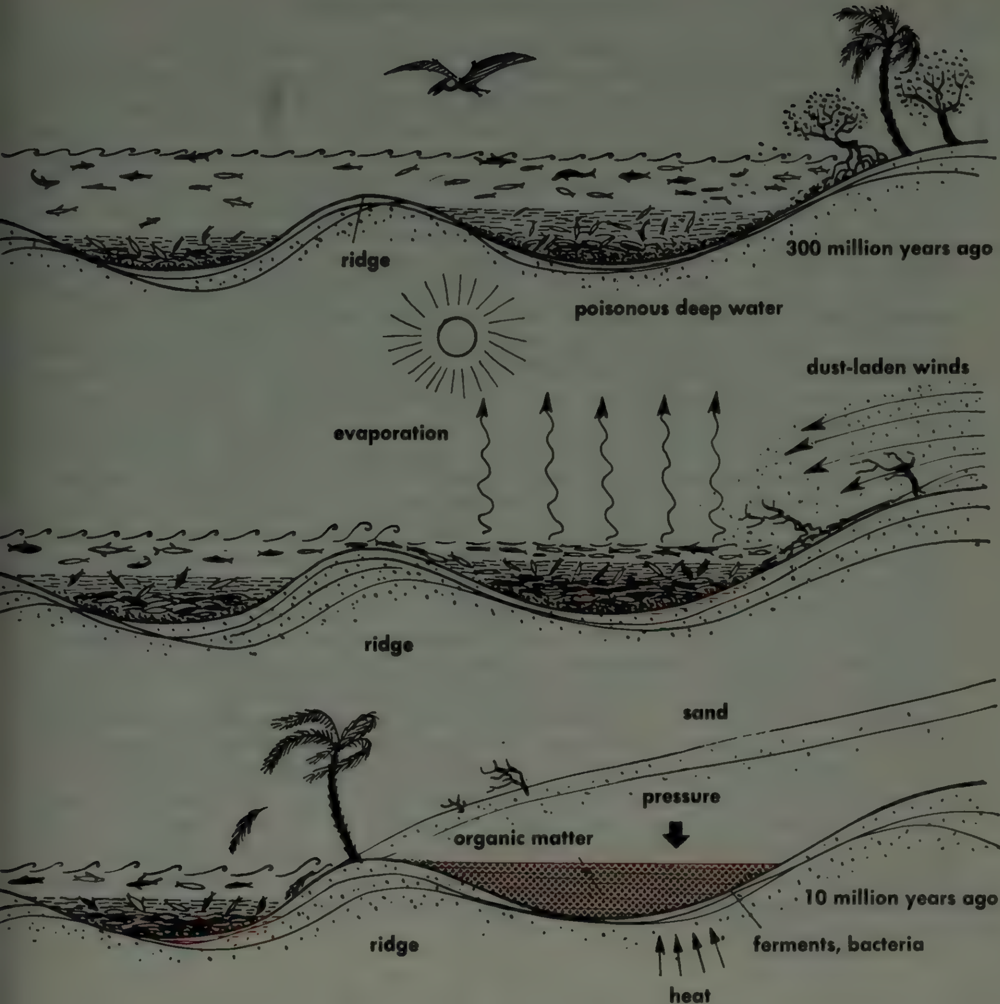
Fig. 4b

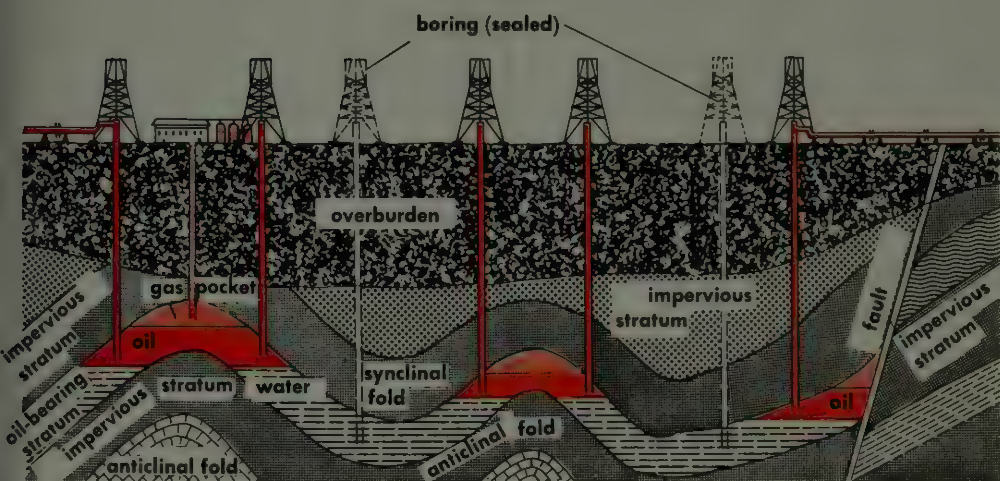
About 300 million years ago the conditions for the subsequent formation of petroleum (mineral oil) were established in shallow coastal waters by the teeming tiny creatures and plants that lived and died in vast numbers. The ooze formed on the bottom by the remains of these organisms was unable to decompose because of lack of oxygen. As a result of climatic changes, these coastal areas became buried under layers of earth, and the organic remains were subjected to high pressures and temperatures over periods of millions of years. The fats, carbohydrates and proteins were thereby subjected to conditions in which they were decomposed and underwent extensive chemical changes. As a result of these changes, a large number of compounds were formed which all enter into the composition of petroleum (alkanes, aromatic compounds, sulphur compounds, etc.). As the conditions of decomposition varied from one region to another, petroleum found in different parts of the world tends to vary considerably in composition. In some places the decomposition was a very intensive process, with the result that natural gas and petroleum particularly rich in aromatic compounds (benzene derivatives) were formed. In addition to pressure and heat, it is certain that catalysts, ferments and especially bacteria have affected the chemical composition of petroleum crude oils.

The world's petroleum reserves are estimates at upward of 40 milliard[1] tons, and production is running at well over 1 milliard tons per year. Unless extensive new deposits are found (under the seas, in desert regions), it seems likely that the world's recoverable reserves of petroleum will be exhausted in something like half a century from now.

A vast variety of products is obtained from petroleum: petrol (gasoline) for aircraft and vehicles, fuel oil for heating and steam-raising, diesel oil, etc. Also, numerous organic chemicals are manufactured from petroleum and are processed into a wide range of products: synthetic rubber, pesticides, synthetic fibres, solvents, drugs, etc. Large-scale exploitation of the world's petroleum reserves started about a hundred years ago and is now a major factor in the power resources available to modern man.

1. A milliard is equivalent to one billion in U.S.A.

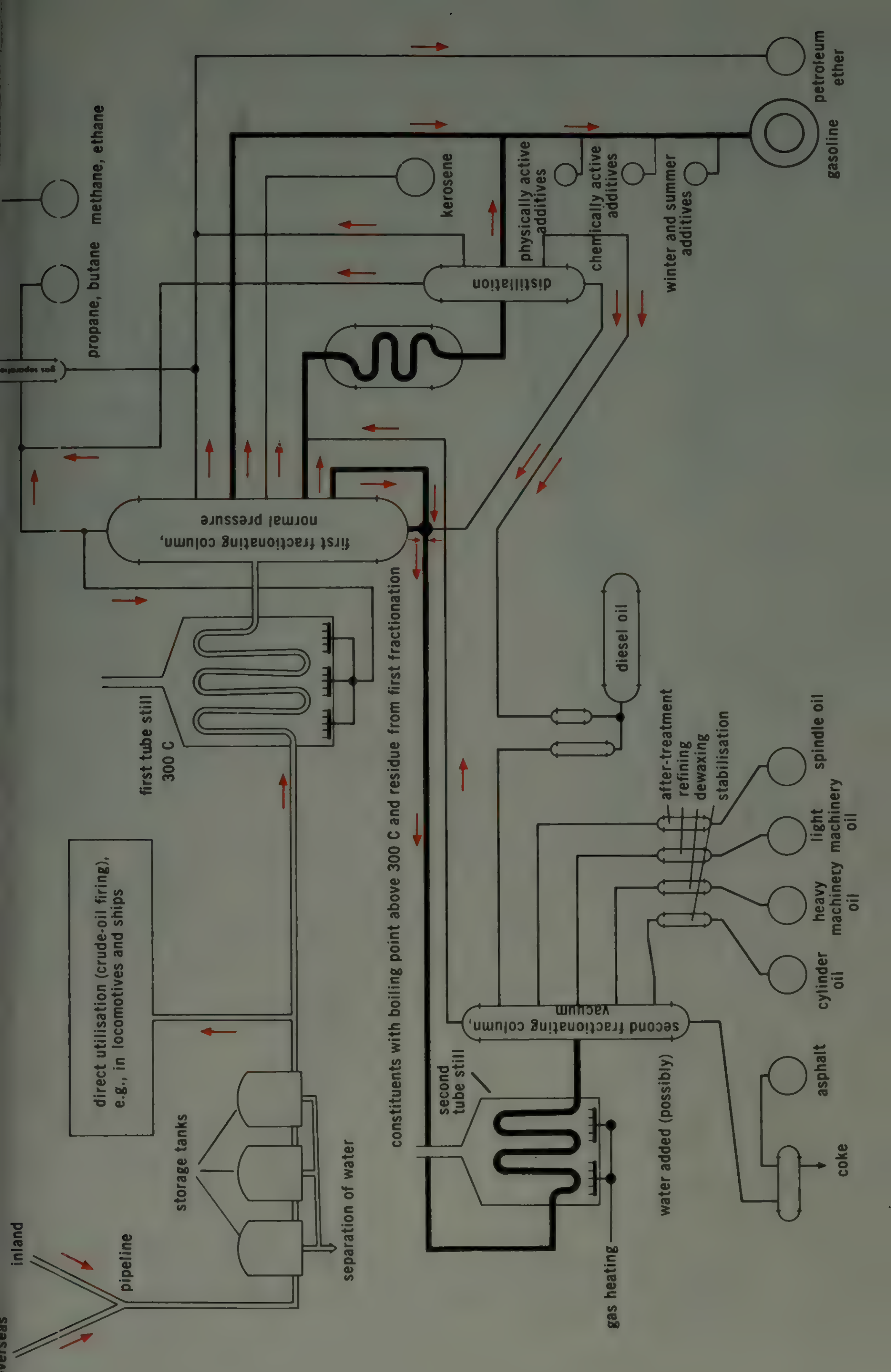




petroleum deposits (schematic)

Petroleum is a complex mixture of hydrocarbons of varying volatility, together with small quantities of substances which contain oxygen, nitrogen, sulphur and ash derived from the vegetable and animal organisms from which the petroleum was formed (see page 34). The crude oil, which is conveyed to the refinery by pipeline, oil tanker or tank wagon, first has water and solid contaminants removed from it by sedimentation and is then split up by fractional distillation (see page 14). The crude oil is pumped through tube stills in which it is heated to 280°–300° C. It is then admitted to a large fractionating column in which the gases, the readily volatile petrol constituents, and the kerosene (paraffin oil) are distilled off. The remaining distillation residue, which is already of a viscous consistency, is pumped through a second tube still, in which it is reheated, and is then passed to a second fractionating column. In this column, which operates under vacuum, various grades of oil are distilled off (gas oil, diesel oil, cylinder oil, machine oil, etc.), while asphalt, mineral pitch, coke-like residues and inorganic matter remain behind. Except for the gas, nearly all the petroleum fractions require further processing whereby their content of deleterious impurities (ash, sulphur and nitrogen compounds, "gumming" and polymerising substances) is reduced or these impurities are removed altogether, either by chemical conversion (treatment with appropriate chemicals) or by physical adsorption with such substances as active charcoal, silica gel, kieselguhr (diatomaceous earth), fuller's earth, etc. Many petroleum fractions have to be treated with additives in order to acquire the desired properties. For example, petrol (gasoline) must undergo further chemical processing to give it good anti-knock and ignition properties, reduce its odour, and make it resistant to ageing. Similarly, machine oils have to be non-resinous, pale-coloured, odourless and oxidation-resistant; additives which further improve the properties of the oil are also employed.

Despite the separation of the multiple mixture of which petroleum consists into a number of fractions, each of which contains fewer constituents than the initial crude petroleum, each fraction still comprises many different constituents (for example, petrol contains upwards of a hundred). The correct fractionation of petroleum is therefore a difficult art which, in addition to the necessary knowledge, involves the use of much complex measuring equipment and costly apparatus. For instance, petroleum contains corrosive substances. Because of this the giant modern petroleum fractionating columns are made of special high-grade steel. Their operation is well-nigh fully automatic.


inland
overseas
pipeline
storage tanks
separation of water
direct utilisation (crude-oil firing), e.g., in locomotives and ships
first tube still 300 C
first fractionating column, normal pressure
gas separator
propane, butane
methane, ethane
kerosene
distillation
physically active additives
chemically active additives
winter and summer additives
gasoline
petroleum ether
constituents with boiling point above 300 C and residue from first fractionation
diesel oil
after-treatment
refining
dewaxing
stabilisation
spindle oil
light machinery oil
heavy machinery oil
cylinder oil
second tube still
second fractionating column, vacuum
water added (possibly)
gas heating
asphalt
coke

Refinery at Cressier, Switzerland
Photo Roland Schneider, Len Sirman Press

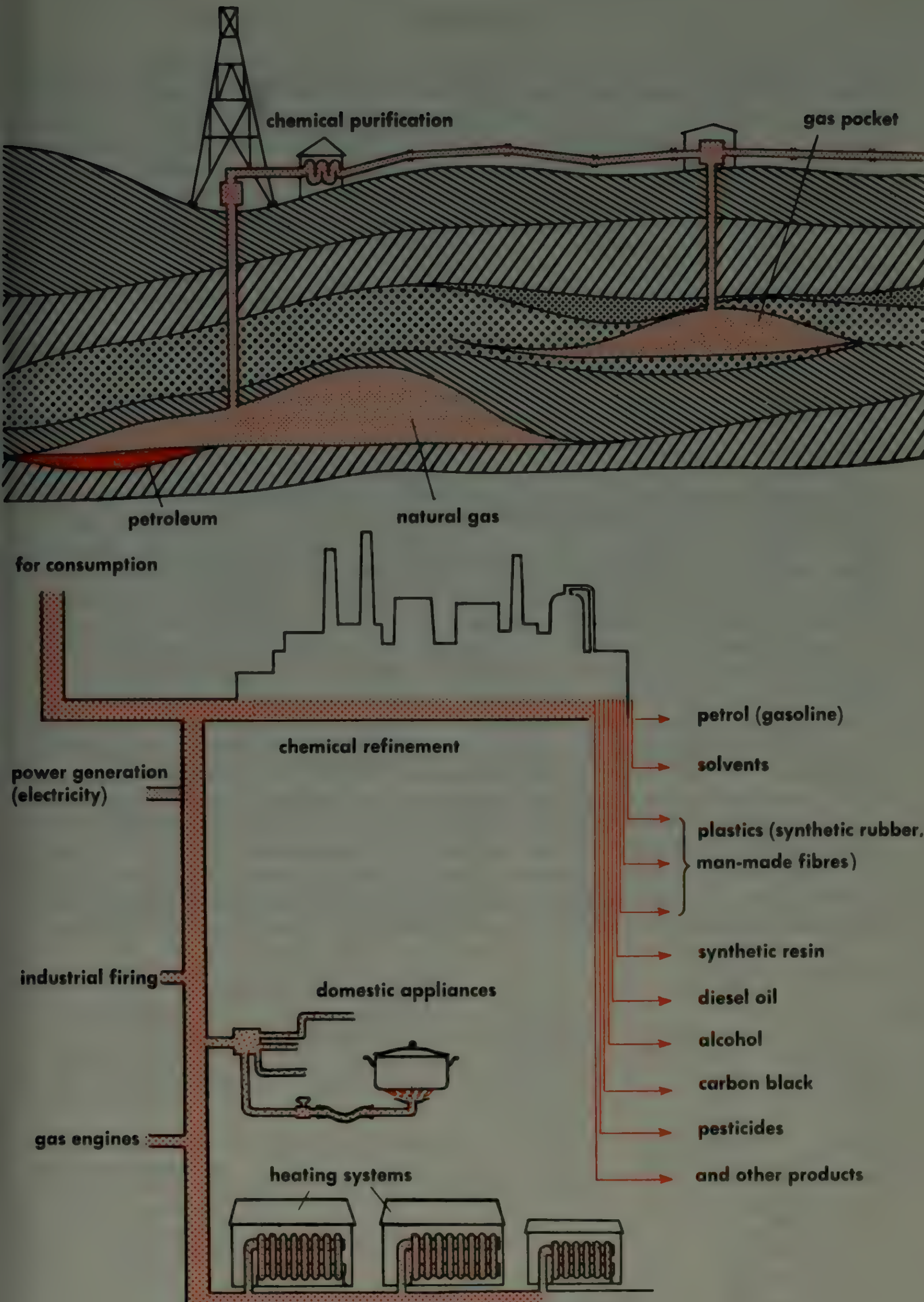
Something like 700 milliard[1] cubic yards of gas a year are produced by gas wells in various parts of the world. Most of this is combustible gas now generally referred to as "natural gas" ("marsh gas" is an older name for very much the same thing); it consists of 80–95% methane, a hydrocarbon.

Some natural gas wells also produce incombustible or highly toxic gases such as carbon dioxide, sulphur dioxide, hydrogen sulphide and others, in rare cases also inert gases (noble gases) such as helium, neon and argon. The sulphurous gases for the most part are of volcanic origin and are formed deep down in the earth's crust. Natural gas in the more specific sense of a combustible hydrogen (largely methane) is usually found in regions where petroleum is also likely to occur. Many gas wells produce so-called "wet" natural gas containing petrol vapour which can be separated and utilised. For example, in the United States, up to 10% of the country's petrol consumption is supplied from this source. In dry natural gas, in addition to methane, small quantities of the following gases are often found to be present: carbon monoxide, hydrogen, helium, neon, argon, and nitrogen. Natural gas provides a very significant part of the raw material for energy production: in terms of energy content, the annual output of natural gas corresponds to 350 million tons of coal. A large proportion of the natural gas is conveyed through huge pipelines from the wells to the industrial centres and major cities, where it is used for industrial and domestic heating. A proportion is also used as fuel for gas engines which drive electricity-generating plant. Also, an increasing proportion of the natural gas output is chemically processed into motor fuels (petrol, diesel oil), plastics, man-made fibres, synthetic rubber, anti-freeze preparations, alcohols, solvents, insecticides, etc.

Besides the naturally existing gas wells, increasing numbers of artificial wells are drilled for tapping the underground gas supplies. Some of these wells deliver gas at pressures of 2000 lb./in^2. and upwards. Sometimes former gas-fields (worked out wells) are used as underground storage reservoirs—sometimes of $1\frac{1}{2}$ milliard cubic yards capacity—for natural gas which may first have been chemically cleaned and had their petrol content removed.

Large quantities of natural gas are found at the foot of the Western Pyrenees. In recent years, too, very large supplies of natural gas have been discovered in Holland.

1. A milliard is equivalent to one billion in U.S.A.


chemical purification
gas pocket
petroleum
natural gas
for consumption
chemical refinement
power generation (electricity)
industrial firing
domestic appliances
gas engines
heating systems
petrol (gasoline)
solvents
plastics (synthetic rubber, man-made fibres)
synthetic resin
diesel oil
alcohol
carbon black
pesticides
and other products

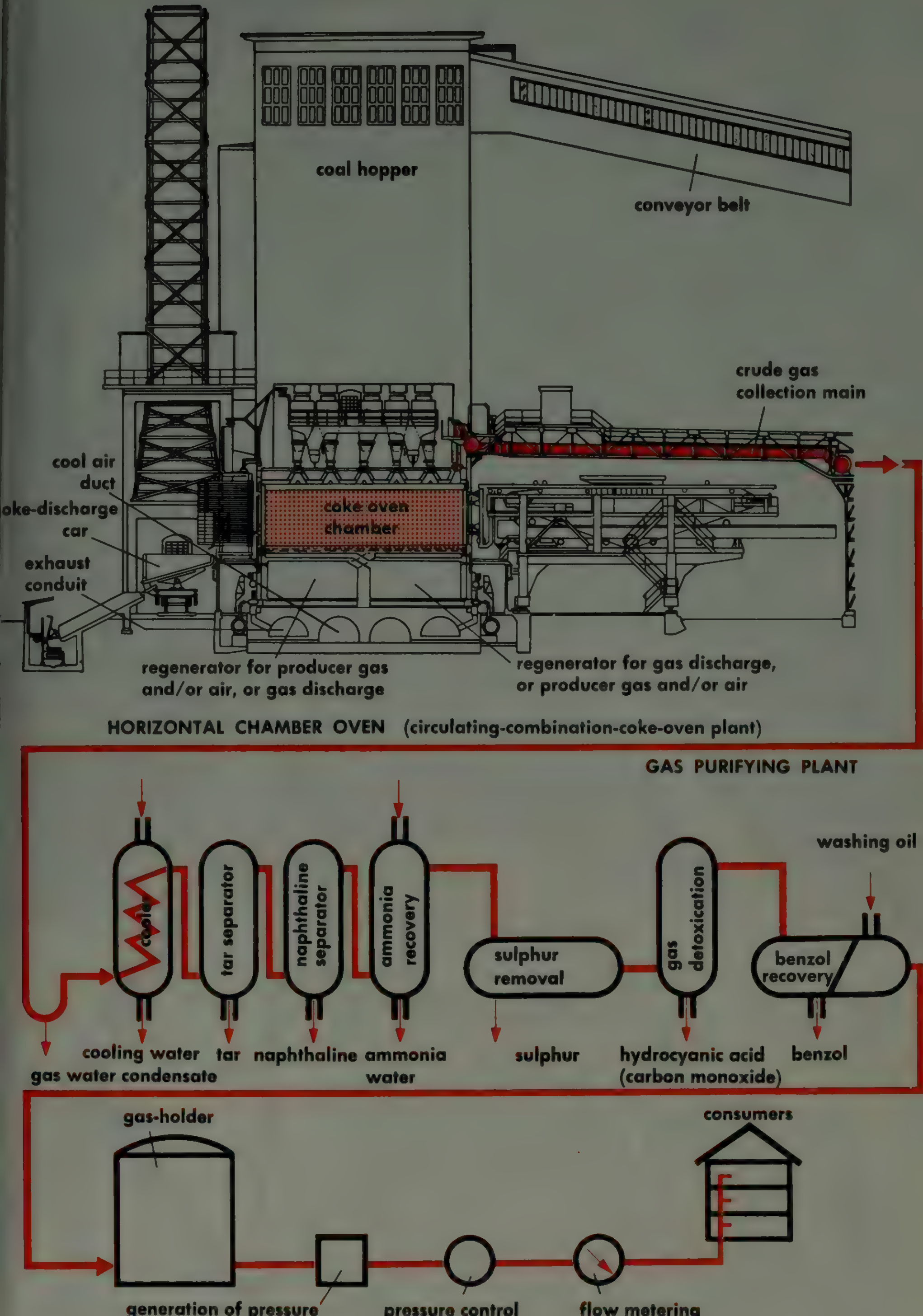
TOWN GAS[1]

Town gas is a combustible and toxic mixture comprising 50% hydrogen, 20–30% methane, 7–17% carbon monoxide, 3% carbon dioxide, 8% nitrogen and 2% hydrocarbons. Furthermore, town gas contains ammonia, sulphur, hydrocyanic acid, benzene and other substances. Sulphur (in the form of strong-smelling chemical compounds) produces the characteristic "gas smell"; indeed, sometimes the smell is deliberately added to provide a warning in the event of the escape of gas. De toxicated gas contains little or no carbon monoxide. It is distributed to the consumers through pipelines at a pressure of at least about 1 lb./in^2. The quality of the gas is rated in terms of its calorific value, i.e., the amount of heat that is produced when one cubic foot of gas is burned with air. The calorific value ranges from 450 to 500 B.T.U. per cubic foot.

Because of its carbon monoxide content, town gas is highly poisonous, and because of its content of combustible gases it is highly explosive when mixed with air. If several cubic yards of gas are allowed to escape into a room, an explosive mixture will be formed, which can be ignited even by a tiny electric spark, such as may, for instance, be caused by a door bell or a telephone ringing.

Town gas is a fuel gas. The term "coal gas" is sometimes applied to it, denoting that it is produced mostly from coal. One way to produce this gas is to heat rough coal to a temperature of 1000°–1200° C, out of contact of air, in a chamber called a retort, which may be of the inclined type. In this process, up to about 500 cu. ft. of town gas can be produced from 100 lb. of dry coal with a low ash content. What remains of the coal after extraction of the gas is called coke. In other gas-making processes the coal is not heated by the external application of heat, but is converted into gas by partial combustion with oxygen and chemical reaction with water vapour. The crude gas produced must be carefully purified: in particular, it is necessary to remove volatile sulphur and nitrogen compounds. In many cases the "water gas" or "producer gas" obtained in this way is used as an admixture to coal gas (up to 40% being added). Fuel gases from oil refineries (which gases are produced by gasification of petroleum) or natural gas are playing an increasingly important part in town gas supply. Despite its dangerous character, town gas is unlikely to be entirely superseded by electricity in the foreseeable future. Gas will be able to hold its own because of its relatively low cost and the convenience with which it can be distributed through pipes. At the same time, the large quantities of carbon dioxide and the by no means inconsiderable quantities of sulphur dioxide, which are formed in the combustion of gas, add to the pollution of the atmosphere of our towns.

1. Coal gas in U.S.A.

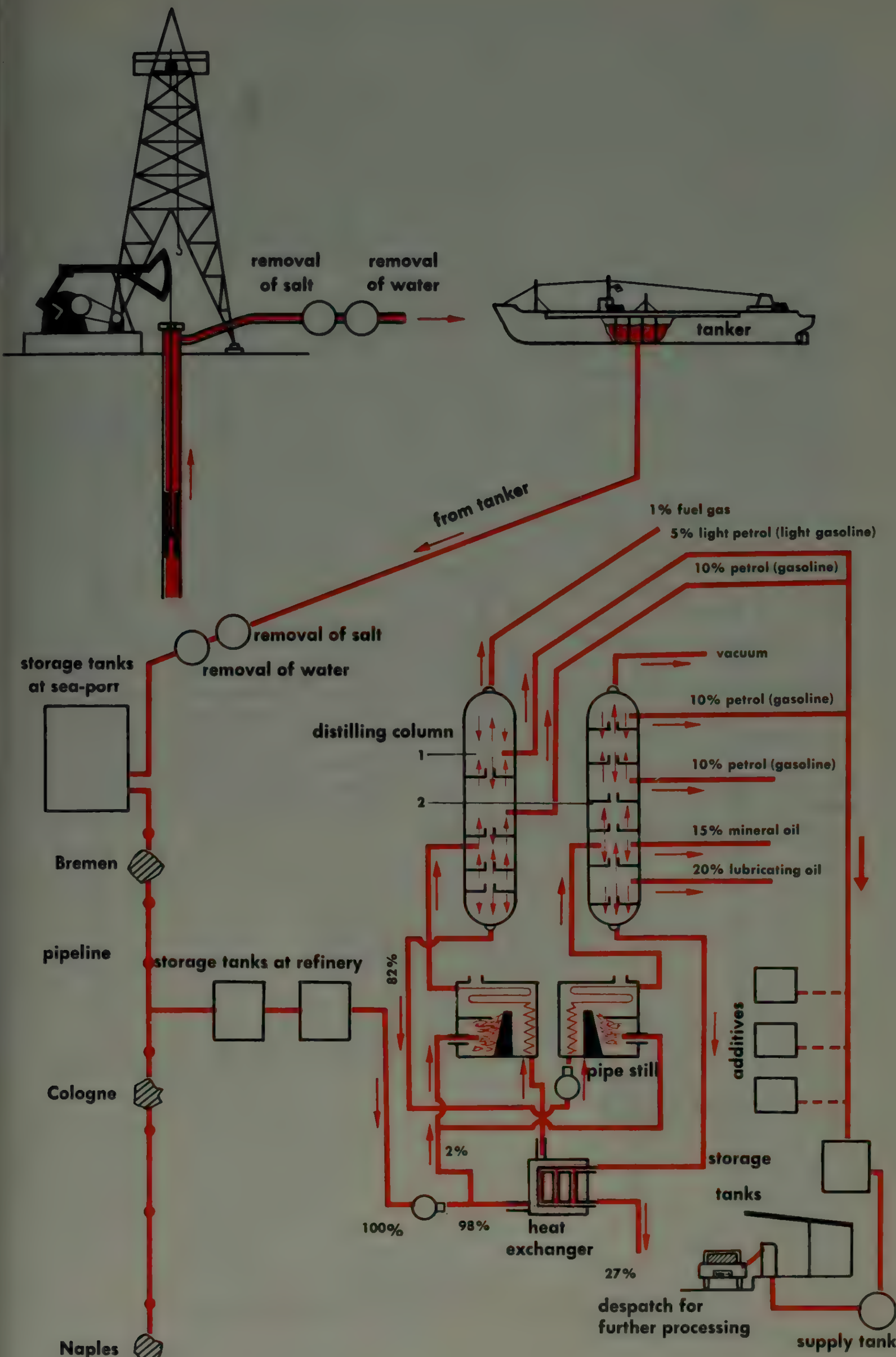

coal hopper
conveyor belt
crude gas collection main
cool air duct
oke-discharge car
exhaust conduit
coke oven chamber
regenerator for producer gas and/or air, or gas discharge
regenerator for gas discharge, or producer gas and/or air
HORIZONTAL CHAMBER OVEN (circulating-combination-coke-oven plant)
GAS PURIFYING PLANT
washing oil
cooler
tar separator
naphthaline separator
ammonia recovery
sulphur removal
gas detoxication
benzol recovery
cooling water
gas water condensate
tar
naphthaline
ammonia water
sulphur
hydrocyanic acid (carbon monoxide)
benzol
gas-holder
consumers
generation of pressure
pressure control
flow metering

Gas-holder, Bienne, Switzerland
Photo Roland Schneider, Len Sirman Press

NATURAL PETROL (GASOLINE)

A distinction is made between synthetic petrol, which is produced from coal and other raw materials by chemical processes, and natural petrol, most of which is obtained as a substance already present in petroleum (mineral oil). The name "petrol" (or "gasoline") denotes a mixture of liquid, volatile hydrocarbons or, to be more precise, a mixture of alkanes, naphthenes and aromatic compounds with boiling points between 40° and 180° C. "Hydrocarbons" is a general designation for chemical compounds which consist solely of the elements hydrogen and carbon and which readily burn to produce carbon dioxide and water if they are mixed with a sufficient quantity of air and then ignited. Petrol for use as a fuel for internal combustion engines is produced by the following process:

The petroleum is pumped from the well through pipelines to storage tanks at the port of shipment, where the crude oil undergoes a preliminary purification treatment. Tankers convey the crude oil to other ports, where it is discharged into storage tanks. From here it is distributed to the refineries, e.g., through pipelines. At the refinery the petroleum is preheated in heat exchangers, then passed to tube stills, where it is heated to a high temperature in special steel tubes. These stills are fired with oil which is likewise obtained from the crude oil. The crude oil, heated to a temperature of several hundred degrees, expands in the distilling column, where it is separated into the fractions: power gas (1% of the total quantity), light petrol (5%) and petrol (10%). The remaining 82% of the original quantity is again passed through the tube still, is reheated to a high temperature, and is passed to a distilling column in which a vacuum is applied, because the distillation temperature can be kept considerably lower when the vacuum is employed. In this second column 20% of the original quantity of crude oil is split up into petrol, 15% into fuel oils and 20% into lubricating oil. The residue, about 27%, provides tar, pitch and coke or undergoes further processing whereby, in some cases, more petrol is produced. However, such petrol is more properly to be regarded as synthetic petrol. The various petrol fractions are mixed and refined; the composition of the mixture depends on the time of year (in the winter the proportion of light petrol in the mixture is higher than in the summer). "Refinement" involves various processing treatments whereby the quality of the petrol as a motor fuel is improved, e.g., it comprises the admixture of aromatic compounds, anti-knock agents, anti-oxidants (ageing inhibitors), etc. The final result must be a volatile fuel which must, among other properties, have a minimum octane number of 80 to 90, ignites easily, does not gasify at room temperature, does not develop "gumming", does not smell objectionably, and burns without residue. Such a mixture of substances is of extremely complex composition, comprising over two hundred individual constituents.

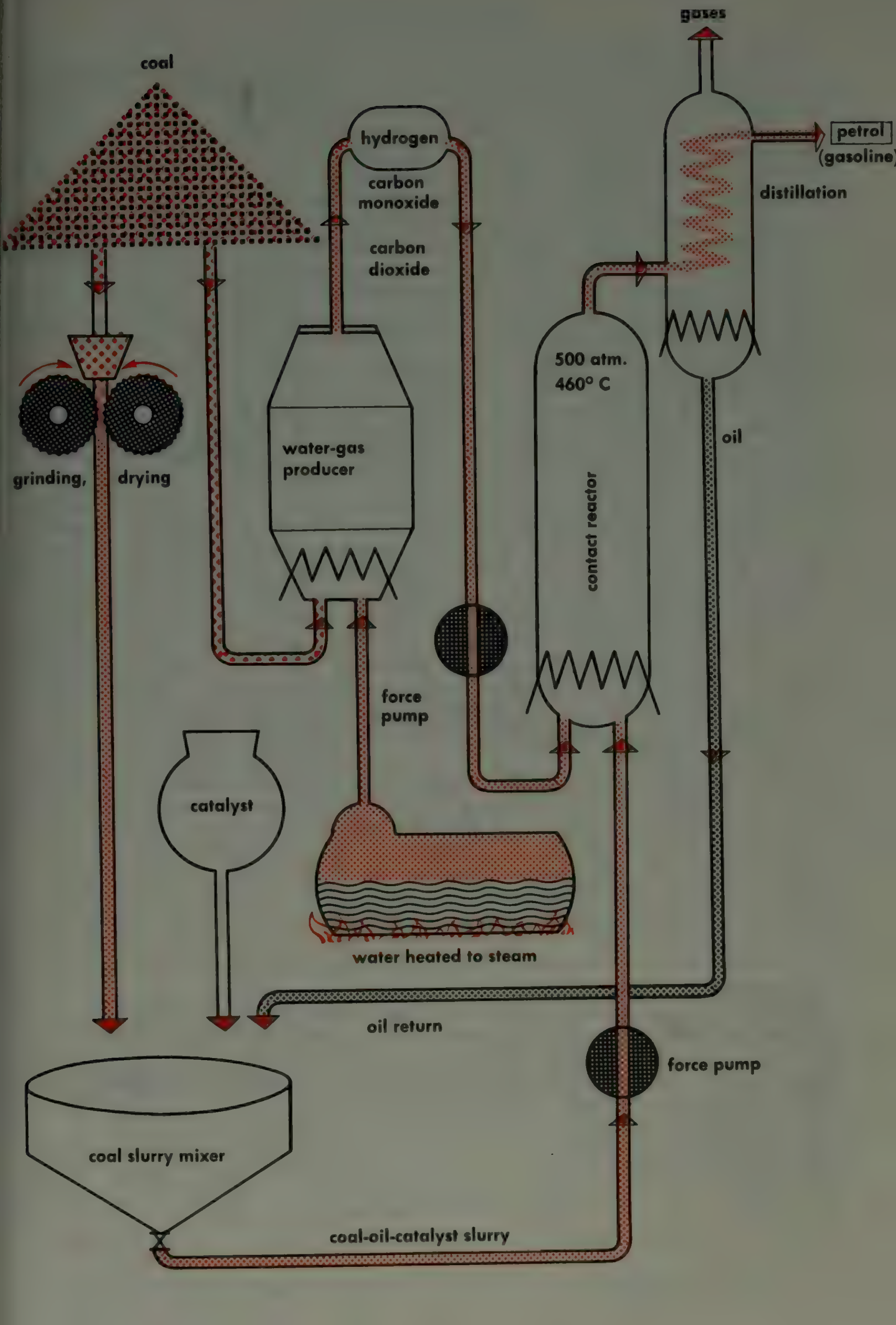

removal
of salt
removal
of water
tanker
from tanker
1% fuel gas
5% light petrol (light gasoline)
10% petrol (gasoline)
removal of salt
removal of water
storage tanks
at sea-port
vacuum
10% petrol (gasoline)
distilling column
1
2
10% petrol (gasoline)
15% mineral oil
20% lubricating oil
Bremen
pipeline
storage tanks at refinery
82%
additives
pipe still
Cologne
2%
storage
tanks
100%
98%
heat
exchanger
27%
despatch for
further processing
supply tank
Naples

SYNTHETIC PETROL (GASOLINE)

In some countries without petroleum resources of their own the production of "synthetic" petrol from coal or from gases containing carbon is of considerable importance. But even in countries which are well supplied with petroleum, a certain amount of synthetic petrol is usually added to the natural product. Synthetic petrol is manufactured either from coal or from natural gas or from petroleum fractions. In the last-mentioned case the designation "synthetic" is appropriate only if the initial substance is gaseous or of low volatility. Crude oil residues and other residues rich in carbon can be converted into petrol. Such petrol, too, can be called synthetic.

Petrol is a high-energy mixture of hydrocarbons consisting of molecules containing 5 to 12 carbon atoms and 12 to 26 hydrogen atoms.

Coal contains only very little hydrogen. To convert it into petrol it is therefore necessary to introduce hydrogen into the compounds. This is done under high pressure and temperature and in the presence of active catalysts. A great deal of petrol used to be produced by so-called catalytic high pressure hydration. In that process it took 4 tons of coal to produce 1 ton of petrol, pressures up to 700 atm. and temperatures of 410°–460° C being employed. The coal was dried, pulverised, and mixed with heavy oil to form a thick slurry. Catalysts were added, and about 70,000 cu. ft. of hydrogen gas per ton of coal was forced in. The hydrogen was produced from coal and water, the carbon monoxide formed in this process being utilised as fuel gas or converted. Synthetic petrol is also manufactured from water and coal by a process in which carbon monoxide and hydrogen are produced from coke, raw brown coal or brown coal briquettes and, after careful cleaning, are passed over catalysts at low pressure. Solid hydrocarbons, in addition to petrol and other products, are formed in this process. In a more recently developed process, gases containing carbon monoxide are conducted, together with water, over suitable catalysts. The resulting reactions produce petrol, as well as acids, alcohols and other substances. Also, petrol is produced from unsaturated hydrocarbon gases with the aid of catalysts. However, since such synthetic petrols are more expensive to produce than petrols from petroleum, in Western Europe synthetic petrol is nowadays of importance only as an additive for natural petrols so as to adjust their properties to meet the exacting requirements of modern internal combustion engines. For this purpose, synthetic petrols having a high octane number (a criterion for the anti-knock properties) are particularly valuable.


gases
coal
hydrogen
carbon monoxide
carbon dioxide
petrol (gasoline)
distillation
500 atm.
460° C
grinding, drying
water-gas producer
oil
contact reactor
force pump
catalyst
water heated to steam
oil return
force pump
coal slurry mixer
coal-oil-catalyst slurry

In industrial plants where large quantities of high-pressure steam are required for a wide variety of purposes the steam is generated in boilers at pressures of about 175–600 lb./in^2. (in power stations: up to about 2400 lb./in^2.).

The various types of boiler differ fundamentally in the method whereby the heat of the furnace and flue gases causes the water to boil. In the elementary form of "boiler", the domestic kettle (Fig. 1), the heat of the flames is applied to the bottom of the receptacle containing the water. In a simple steam-generating boiler this same principle is applied, and in so-called fire tube boilers the hot furnace gases pass through tubes in the water space (fire tubes). The advantage of this type of boiler is its simple operation. On the other hand, their steam-generating efficiency is relatively low because of the limited grate area and the slow water circulation. Also, it has the disadvantages of having a low operating pressure and taking up much space. For these reasons such boilers are nowadays used only in locomotives or in installations which require fairly small quantities of steam.

Modern high-efficiency boilers, capable of very quickly coping with high peak loads, are of the water-tube type. In such boilers the water is evaporated in tubes which are arranged inside a heated chamber in which they are exposed to the radiant heat of the flames and the hot flue gases. They are constructed as boilers with steeply inclined water tubes (e.g., Stirling boiler, Fig. 2) or with tubes set at a relatively low inclination (e.g., Babcock and Wilcox boiler, Fig. 3). The tubes are interconnected at their ends by so-called headers, which are usually set at right angles to the tubes. The boiler feed-water, preheated by the flue gases, enters the upper steam drum from where it flows through unheated or only slightly heated tubes to the lower headers or drums. From here it ascends into the water tubes, in which it is evaporated and is returned in the form of a water-and-steam mixture to the upper drum. In the latter the steam is separated from the entrained water and flows through the superheater tubes, which are heated by flue gases of sufficiently high temperature. The steam then flows to the consumer equipment. The separated water, together with additional water, flows back to the lower headers or drums, and the cycle is repeated.

The boilers referred to here are heated with coal or oil. A coal-fired boiler is usually provided with a so-called mechanical stoker, a frequently employed form of which is the chain grate (Fig. 4). This device consists of a slowly moving endless chain of grate bars. The coal fed on to one end of the grate is burned in the furnace. The residual matter, slag and ash fall off the other end of the chain into the ashpit. To control the combustion process, air is blown from below through the grate. In other types of firing system pulverised coal, together with air, is blown into the combustion chamber of the furnace, where it burns at temperatures of about 1800° C. Again, in other systems finely divided ("atomised") fuel oil is sprayed through nozzles into the combustion chamber. The latter is lined with refractory material (fireclay brick), in which water pipes are embedded. These absorb the radiant heat, protect the lining and, in addition, produce steam.

Special boilers can generate high-pressure steam (150–350 lb./in^2.). In such boilers the water is circulated by pumping (forced-circulation boilers, Fig. 5), whereby the quantity of water circulated is six or seven times as large as that corresponding to the evaporative capacity. Alternatively, the water is pumped straight through the boiler tubes ("once-through" or flash boiler, Fig. 6).

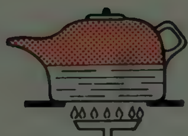


Fig. 1 KETTLE

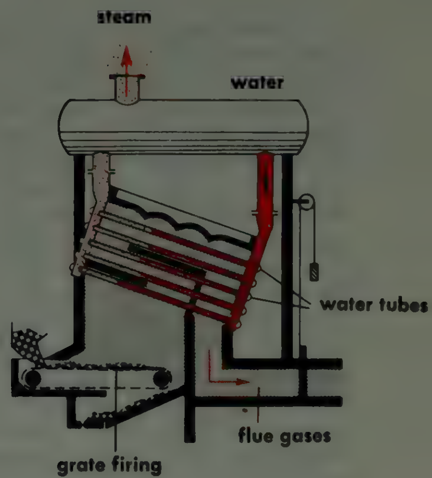


Fig. 3 BABCOCK & WILCOX BOILER

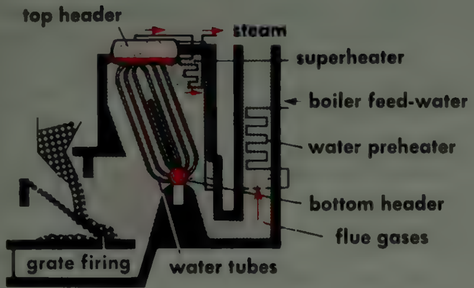


Fig. 2 STIRLING BOILER

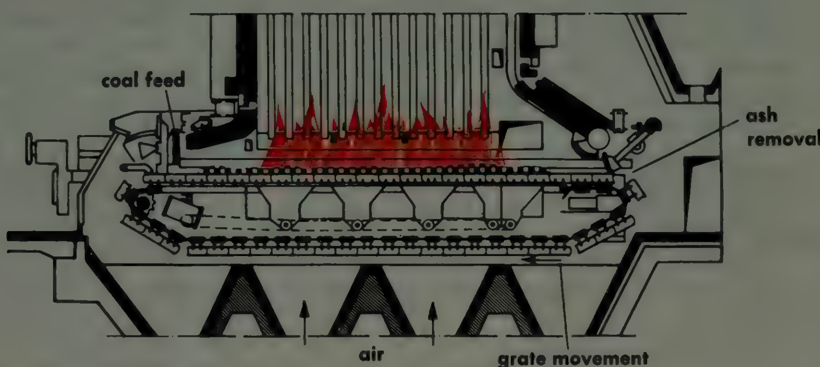


Fig. 4 CHAIN GRATE FIRING SYSTEM

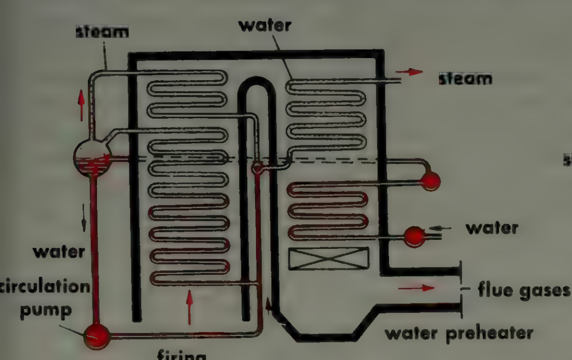


Fig. 5 FORCED-CIRCULATION BOILER

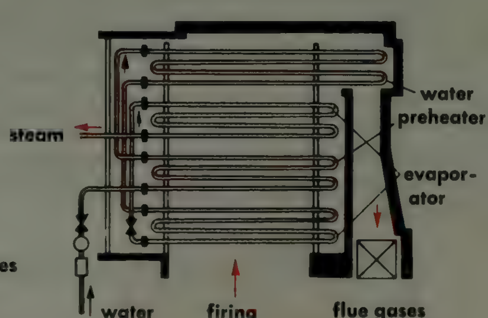


Fig. 6 ONCE-THROUGH BOILER

A steam engine utilises the energy contained in steam under high pressure. The energy that is released when steam expands is made to produce rotary motion which can be used for the driving of machinery. The steam from the boiler is admitted into the cylinder in which there is a piston and in which the steam expands, causing the piston to move (Fig. 1a). When the piston has travelled to the end of the cylinder and thus completed its stroke (Fig. 1b), the now expanded steam is allowed to escape from the cylinder. At the same time the steam is changed over, live steam under pressure being admitted to the other side of the piston, causing the latter to travel back, past its starting point (Fig. 1c), until it has reached the other end of its stroke (Fig. 1d). A steam engine of this kind is called "double-acting" because the force of the steam is applied alternately on two sides of the piston. While the piston is being forced in one direction by the expanding steam, the spent steam is pushed out of the cylinder on the other side of the piston. Reversing, i.e., the change-over of the steam supply so as to ensure the admission of live steam to the appropriate side of the piston and the discharge of the spent steam on the other side, is effected automatically by a control device called a slide valve. On some steam engines, valves are similar in principle to those used in internal combustion engines. The commonest type is the flat slide valve (often called a "D-slide valve") (Fig. 2). It alternately covers the steam inlet port and the exhaust port. During the piston movement the slide valve opens the exhaust port for the escape of the spent steam behind the piston (Figs. 2a and 2b). The slide valve must therefore always be in such a position that it connects the working side of the piston to the live steam supply, and opens the exhaust port on the other side to enable the steam to escape into the exhaust channel. The valve does this by moving to-and-fro at the same rate as the piston. It is controlled through a linkage system from the crankshaft in such a manner that the valve moves in the opposite direction to the piston (Fig. 3). The spent steam that emerges from the cylinder is passed to a condenser where it is cooled and thereby precipitated as water. In some cases (e.g., steam locomotives) it is, instead, discharged direct into the atmosphere.

The term "compound steam engine" refers to an engine in which the steam is expanded in several cylinders (usually three), in successive stages. As a rule, all three pistons have the same length of stroke, since they drive the same crankshaft. However, as the volume of the steam increases as the result of expansion, i.e., when its pressure is reduced, the second (or medium-pressure) stage has a larger piston diameter than the first stage (the high-pressure stage). The third (or low-pressure) stage has a piston of still larger diameter than the second stage. The transmission of power and motion from the piston to the crankshaft is effected through a crosshead (a reciprocating block sliding between guides) which forms the junction piece between the piston-rod and the connecting-rod. One end of the latter is pivotably connected to the crosshead and the other end is connected to the circumference of a wheel or to a crank of the crankshaft. When the piston and, with it, the piston rod and crosshead move to-and fro, the crosshead transmits this movement to the crankshaft or wheel and thus produces rotation. One end of the crankshaft is provided with a flywheel to ensure smooth running, free from jerkiness due to the reciprocating motion of the piston. Nowadays steam engines are used only in cases where slow-running machinery with varying power requirements has to be driven, e.g., for winding gear in coal mines, ships' propulsion machinery, rollers, etc. Even the best steam engines have an efficiency of only 15–18% in terms of utilisation of the energy contained in the coal.

Fig. 1 MODE OF OPERATION OF A STEAM ENGINE

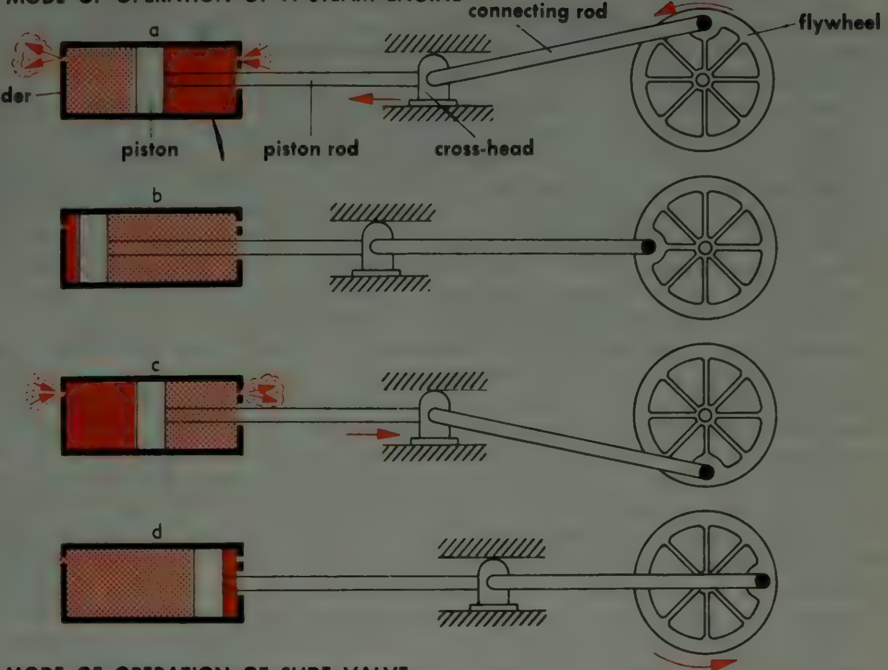


Fig. 2 MODE OF OPERATION OF SLIDE VALVE

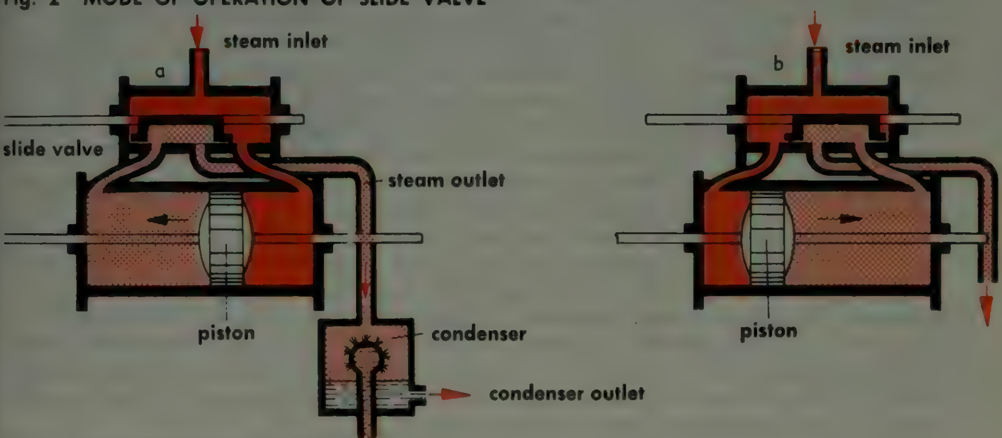
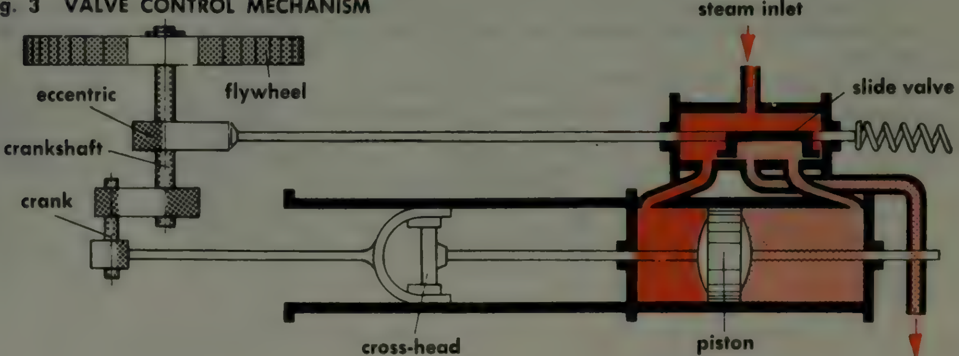


Fig. 3 VALVE CONTROL MECHANISM



In a steam turbine the energy in steam under pressure is used for producing a mechanical rotary motion which is usually employed for electric power generation. The steam is discharged in the form of a high-velocity jet from a nozzle and impinges upon blades mounted on a wheel, whereby the latter is caused to rotate. In order to convert the pressure energy of the steam as efficiently as possible into kinetic energy, so-called Laval nozzles are employed comprising an inlet portion, a constricted throat and a gradually widening outlet (Fig. 1). As a result of the constriction of the passage followed by the widened outlet portion (diffuser), the pressure of the steam flowing through this nozzle is converted into velocity. The higher the velocity of the steam is, the larger is the force that the jet of steam is able to exert upon any obstacle it encounters. Also, it very much depends upon how much the steam is thereby deflected from its original path. If a wheel is provided with a set of blades which deflect the steam into very nearly the circumferential direction (Fig. 2), then the steam will exert upon the wheel a force in the circumferential direction, causing it to rotate (Fig. 3). In order to achieve the fullest possible utilisation of the energy contained in the steam, a number of successive stages are usually arranged one behind the other. The wheels provided with blading are mounted on the same shaft and therefore all revolve at the same speed. The rotating assembly as a whole is called the rotor. The decrease in pressure and the increase in diameter of the successive blade wheels is effected in definite and carefully calculated stages. Theoretically the entire pressure of the steam (e.g., 1500 lb./in.^{2}) can be expanded to atmospheric pressure in a single nozzle and be converted into velocity. In that case the blade wheel would have to revolve at a very high circumferential velocity in order to attain reasonable efficiency. Such a high velocity would, however, present a number of technical difficulties and involve the risk of destruction of the machine by the large centrifugal forces that would arise. Turbines embodying the utilisation of the energy in successive stages are designed according to various alternative principles:

1. Velocity staging: In this arrangement the entire steam pressure is, indeed, converted into velocity in one nozzle, but this velocity is utilised in stages in a number of blade rings (rows of moving blades mounted on the same shaft (Curtis turbine, Fig. 4).
2. Pressure staging: The steam is expanded in a nozzle, and the resulting velocity is used up in driving a blade ring mounted behind the nozzle. The steam then passes through guide blading which functions as a nozzle and in which the steam is again expanded a little and the velocity used for driving another blade ring (Fig. 2). This alternating sequence of stationary nozzles and rotating blades is continued until the entire steam pressure has been used up. The same quantity of steam (in terms of weight) flows successively through all the blade rings that make up the turbine rotor. However, the stage-by-stage reduction of the high initial pressure of the steam is associated with a progressive increase in volume, and for this reason the flow sections between the blades of the successive rows of blading must become correspondingly larger. For this reason the blade rings at the end of the turbine are of larger outside diameter than those at the beginning (impulse turbines).
3. Reaction staging: In this system the pressure is converted into velocity not only in nozzles or guide blades, but also in the blade rings of the rotor (reaction turbines).

Steam turbines can also be classified according to the utilisation of the steam pressure drop. For instance, a condensing turbine (Fig. 5) utilises the entire available drop from high pressure to the vacuum in the condenser; a back-pressure turbine utilises only the top part, whereas an exhaust-steam turbine utilises only the bottom part of the pressure drop.

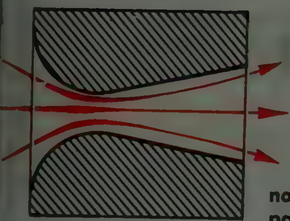


Fig. 1 LAVAL NOZZLE (*schematic*)

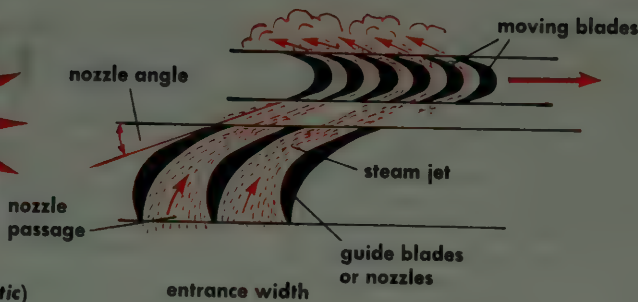


Fig. 2 DEFLECTION OF THE STEAM JET (*pressure staging*)

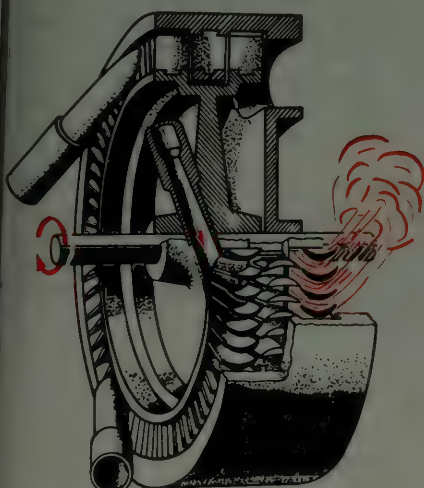


Fig. 3 DRIVE OF A WHEEL WITH ROTOR BLADES (*schematic*)

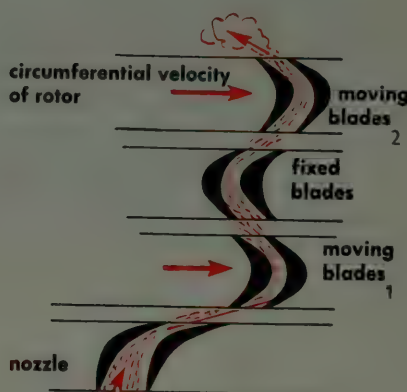


Fig. 4 DOUBLE-ROW CURTIS TURBINE (*velocity staging*)

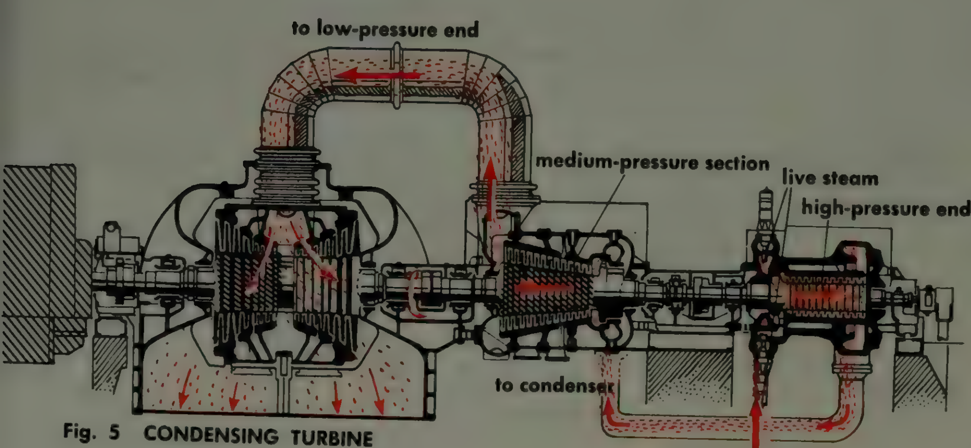


Fig. 5 CONDENSING TURBINE

Gas turbines are driven by the combustion gases from liquid fuels. In construction and operation they resemble steam turbines (see page 44) in that a flowing medium with a high energy content—i.e., the combustion gases—produces a rotary motion as a result of being deflected by rings of blading mounted on a rotor. The operation of a gas turbine is shown schematically in Fig. 1: the compressor draws in fresh air and compresses it to a pressure of 50–75 lb./in.^{2}; the air is forced by the compressor through a heat exchanger (the regenerator) where it is preheated by the heat that is still present in the exhaust combustion gases emerging from the turbine; and finally the preheated air is admitted into the combustion chamber. In this chamber liquid fuel is burned, thereby producing gases with a temperature of about 650° C. These combustion gases flow at high velocity into the turbine and drive it.

The turbine itself, the compressor, and the electric generator are all mounted on one shaft. The turbine cannot transmit its entire power output to the generator, for a substantial proportion is needed for driving the compressor. The turbine is started with the aid of an electric motor which first has to set the compressor in motion, in order to produce compressed air and supply it to the combustion chamber so as to enable the combustion gases to be formed. Only then can the turbine start running. Fig. 2 shows the main features of a gas turbine. The most familiar form of gas turbine is the jet propulsion engine for aircraft (p. 292 vol. II). This is a particular specialised form of turbine, however. Ordinary gas turbines are used only in cases where the liquid fuels they use are cheaper than coal.

Another type of gas turbine is the hot-air turbine (Fig. 3). Here the rotary is produced, not by combustion gases, but by air which is heated in a heat exchanger by the hot combustion gases. This hot air recirculates continuously, i.e., it is drawn into the compressor again after leaving the turbine. However, the air must be cooled before entering the compressor, otherwise the blading of the latter would soon be destroyed as a result of operating at excessively high temperatures.

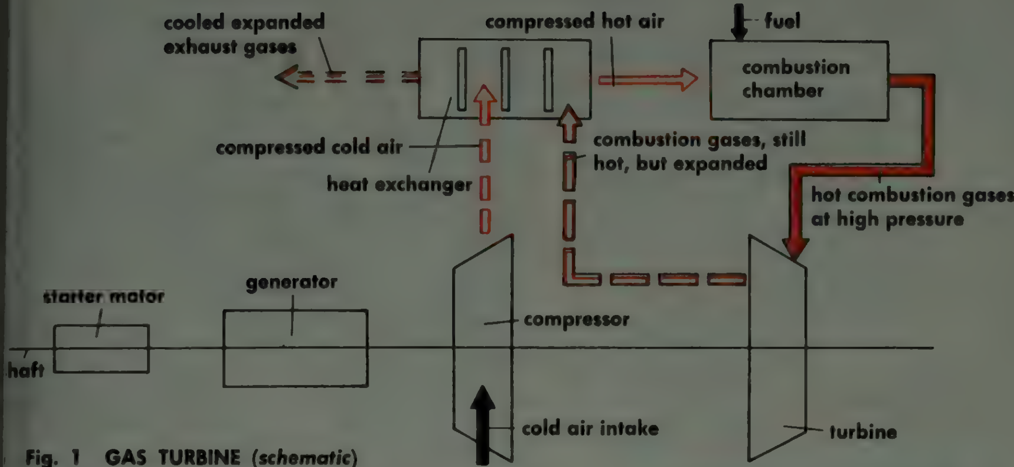


Fig. 1 GAS TURBINE (*schematic*)

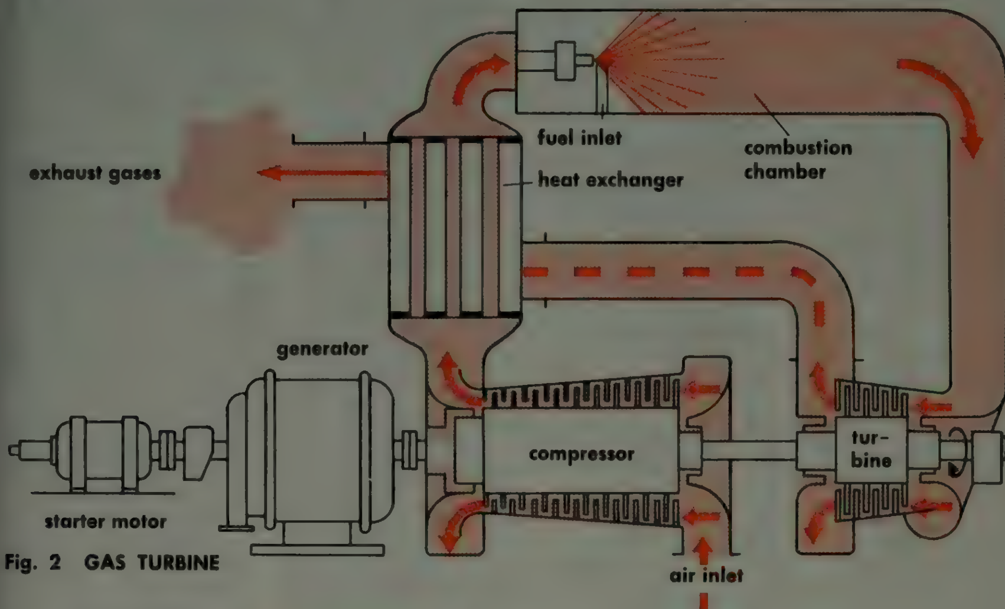


Fig. 2 GAS TURBINE

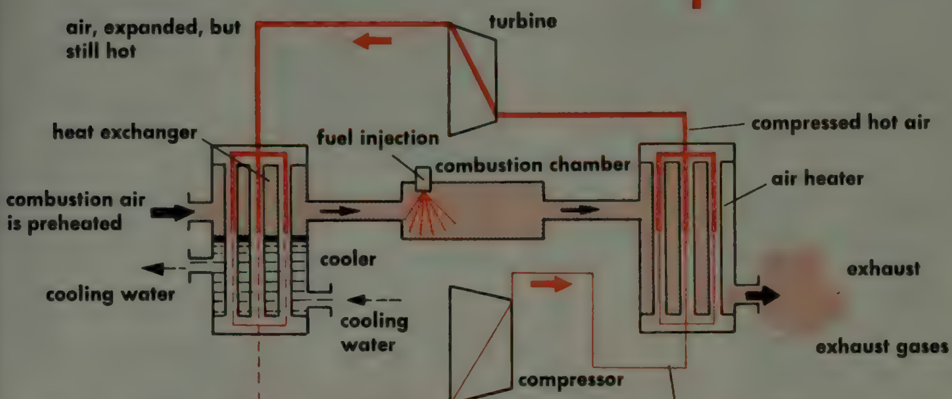
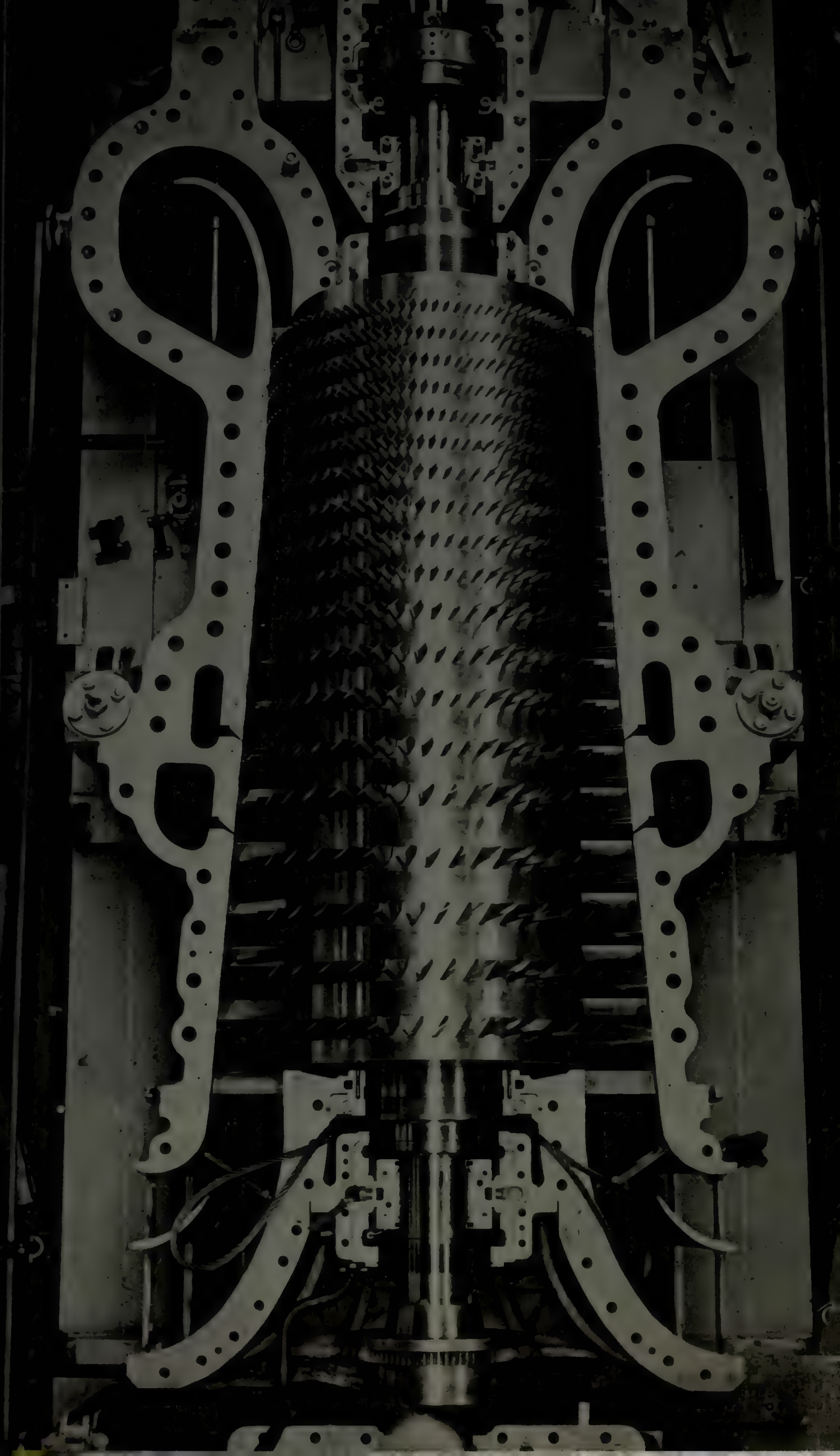


Fig. 3 HOT-AIR TURBINE

Gas-turbine, built by Brown-Boveri, Baden, Switzerland
Photo Roland Schneider, Len Sirman Press



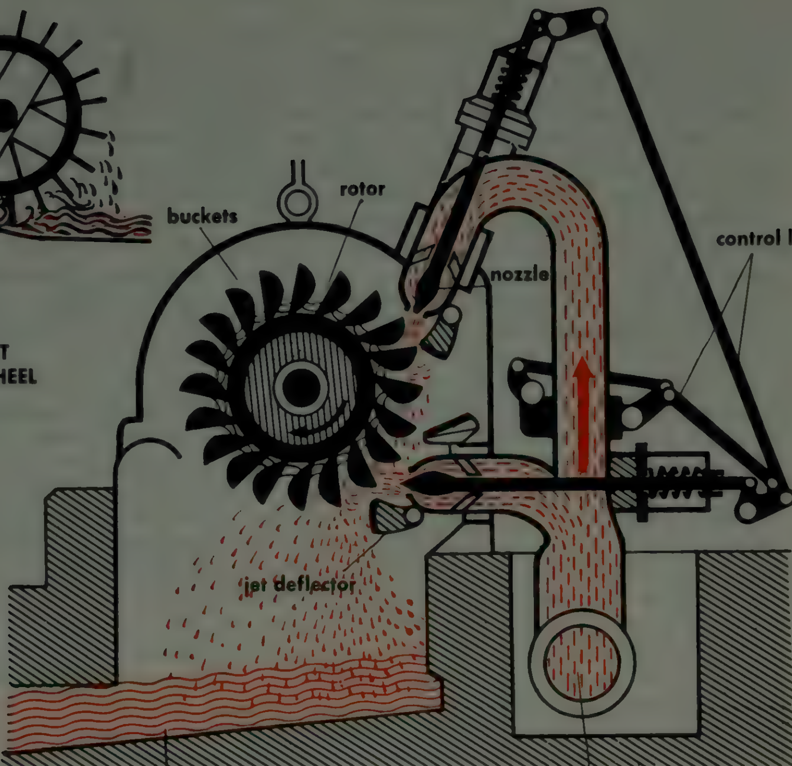
WATER TURBINES (PELTON)

In water turbines the kinetic energy of flowing or falling water is converted into mechanical rotary motion.

The oldest form of "water turbine" is the water wheel. The natural head—difference in water level—of a stream is utilised to drive it. In its conventional form the water wheel is made of wood and is provided with buckets or vanes round the periphery. The water thrusts against these, causing the wheel to rotate. The latter drives the millstones (or sometimes other machinery). In the case of an "overshot wheel" the water pours on to the buckets from above. If the water thrusts against the vanes on the underside of the water wheel, as in Fig. 1, it is called an "undershot wheel". The principle of the old water wheel is embodied in the modern Pelton wheel (Fig. 2), which consists of a wheel provided with spoon-shaped buckets round the periphery (Fig. 3). A high-velocity jet of water emerging from a nozzle impinges on the buckets and sets the wheel in motion. The speed of rotation is determined by the flow rate and the velocity of the water; it is controlled by means of a needle in the nozzle (the turbine operates most efficiently when the wheel rotates at half the velocity of the jet). If the load on the wheel suddenly decreases, the jet deflectors partially divert the jet issuing from the nozzle until the jet needle has appropriately reduced the flow (Fig. 4). This arrangement is necessary because if, in the event of sudden load decrease, the jet needle were suddenly closed, the flow of water would be reduced too abruptly, causing harmful "water hammer" phenomena in the water system. In most cases the control of the deflector is linked to an electric generator.

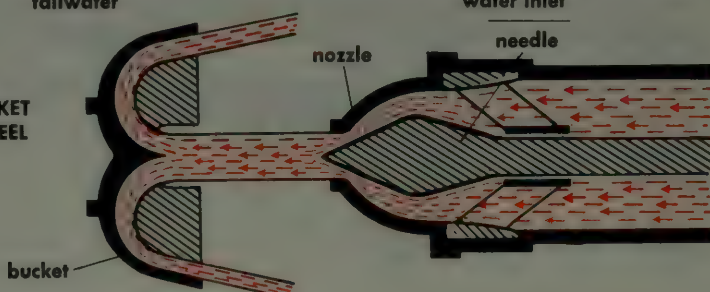
A Pelton wheel is used in cases where large heads of water are available. The water is discharged from the high-level reservoir through pipes to the turbine.

. 1 UNDERSHOT WATER WHEEL



g. 2 TWO-NOZZLE PELTON WHEEL

g. 3 NOZZLE AND BUCKET OF A PELTON WHEEL



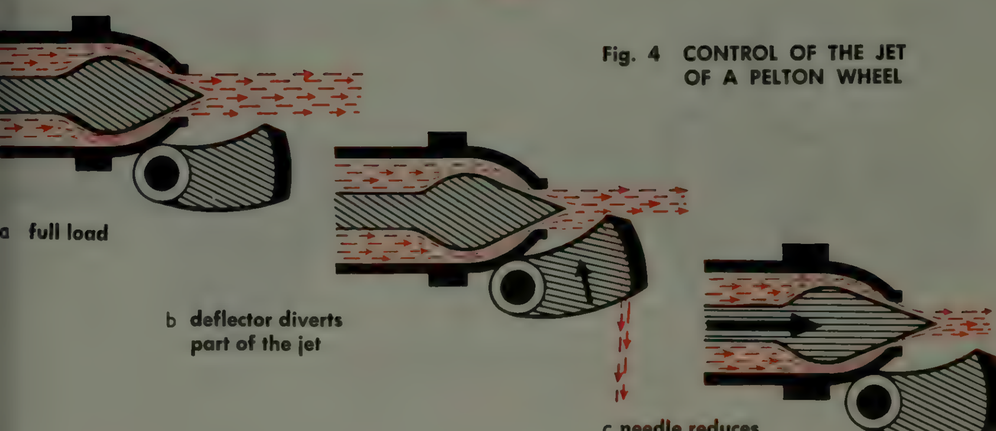


Fig. 4 CONTROL OF THE JET OF A PELTON WHEEL

a full load

b deflector diverts part of the jet

c needle reduces water flow

In the great majority of cases (large and small water flow rates and heads) the type of turbine employed is the *Francis* or *radial-flow turbine*. The significant difference in relation to the Pelton wheel is the water is diverted inside the turbine. This diversion takes place at right angles to the direction of entry (Fig. 1), causing the runner—the turbine rotor—to spin round. The water first enters the volute, which is an annular channel surrounding the runner, and then flows between the fixed guide vanes, which give the water the optimum direction of flow. It then enters the runner and flows radially through the latter, i.e., towards the centre. The runner is provided with curved vanes upon which the water impinges. The guide vanes are so arranged that the energy of the water is largely converted into rotary motion and is not consumed by eddies and other undesirable flow phenomena causing energy losses. The guide vanes are usually adjustable so as to provide a degree of adaptability to variations in the water flow rate and in the load of the turbine.

The guide vanes in the Francis turbine are the elements that direct the flow of the water, just as the nozzle of the Pelton wheel does. The water is discharged through an outlet from the centre of the turbine. The main features of a water turbine of this type are illustrated schematically in Fig. 1. The volute, guide vanes and runner are shown in Fig. 1a. The diversion of the water at right angles to its direction of entry is clearly indicated in Fig. 1b, which is a cross-section through the turbine.

For very low heads and high flow rates—e.g., at barrages[1] in rivers—a different type of turbine, the *Kaplan* or *propeller turbine* is usually employed. It is rather like the propeller of a ship operating in reverse: the ship's propeller rotates and thrusts the water away behind it, thus causing the ship to move forward (p.276, vol. II), in the Kaplan turbine the water flows through the propeller and sets the latter in rotation. The water enters the turbine laterally (Fig. 2), is deflected by the guide vanes, and flows axially through the propeller. For this reason, these machines are referred to as axial-flow turbines. The flow rate of the water through the turbine can be controlled by varying the distance between the guide vanes; the pitch of the propeller blades must then also be appropriately adjusted (Fig. 3). Each setting of the guide vanes corresponds to one particular setting of the propeller blades in order to obtain high efficiency.

The runner of a water turbine drives a shaft which is coupled, directly or through gearing, to an electric generator (see page 78).

1. Dams in U.S.A.

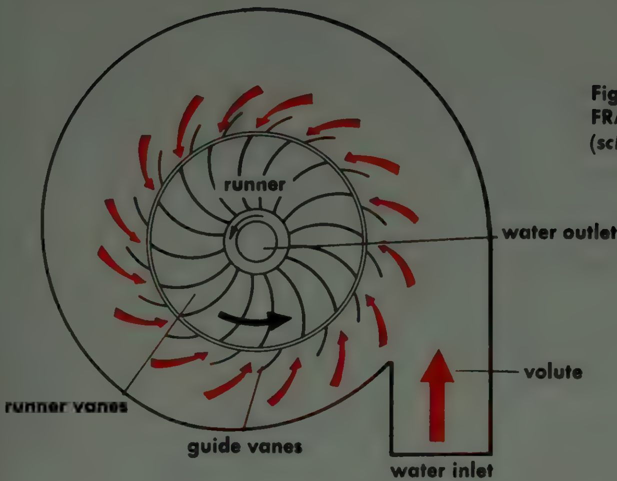


Fig. 1a
FRANCIS TURBINE, SIDE VIEW
(*schematic*)

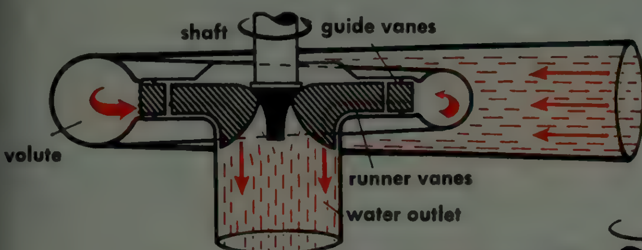


Fig. 1b SECTION THROUGH A FRANCIS TURBINE

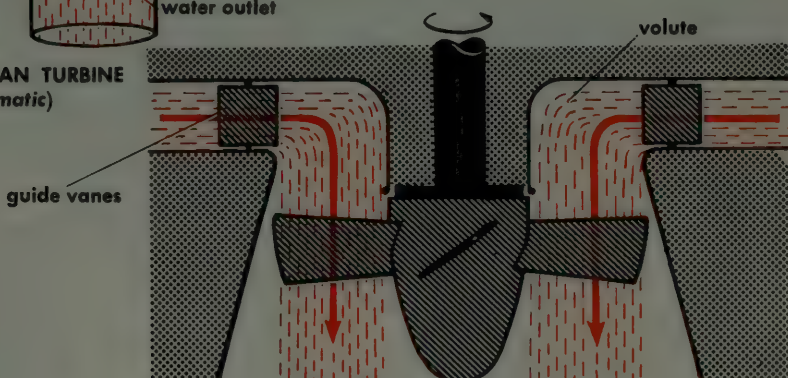


Fig. 2 KAPLAN TURBINE
(*schematic*)

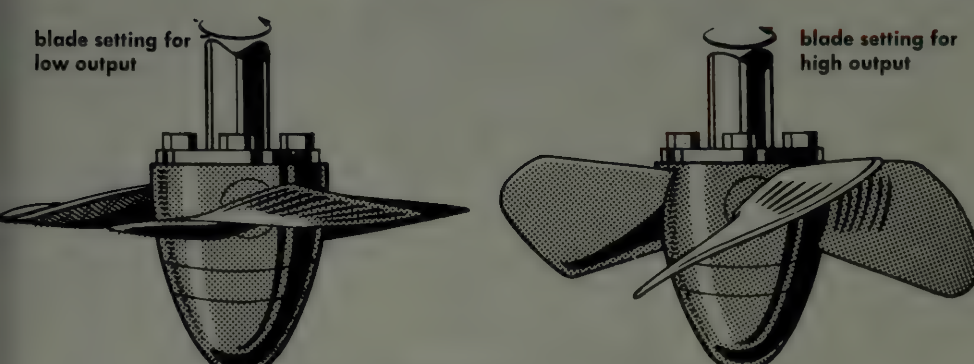
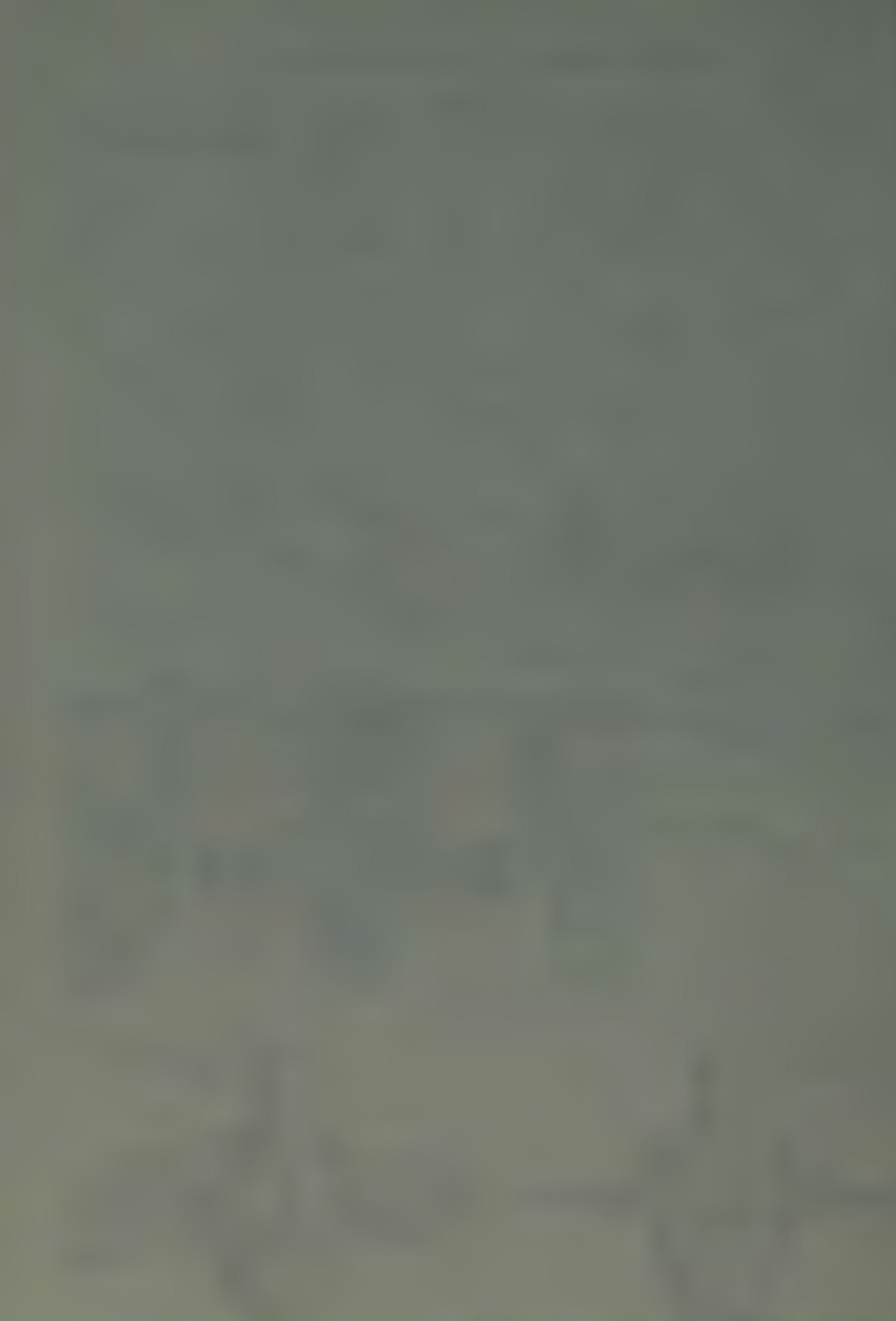
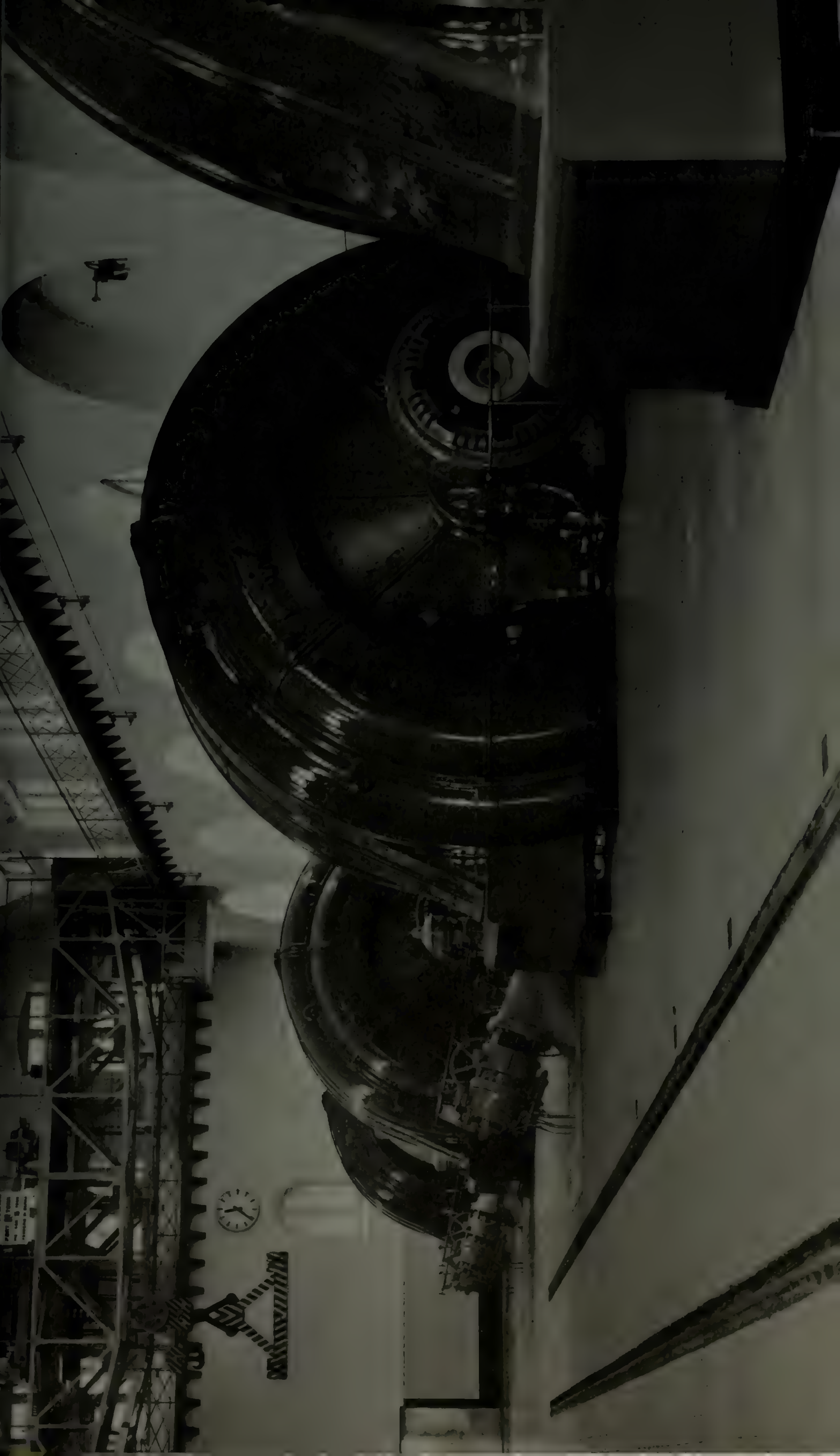


Fig. 3 PROPELLER OF KAPLAN TURBINE



Water turbine in Swiss hydro-electricity station
Photo CFF



A fuel cell is a device in which the energy released in the oxidation of a conventional fuel is made directly available in the form of an electric current. It thus avoids the wasteful detour of the conventional thermal power stations, i.e., the generation of electricity via the "inferior" thermal energy. Although the principle of the fuel cell was formulated by W. Ostwald as long ago as 1894, it is only in recent years that some success has been achieved in the construction of efficient cells of this kind. In the fuel cell constructed by Baurand Ehrenberg in 1911 (Fig. 1) a carbon rod serves as the fuel. It functions as the anode, introducing C^{++++} ions into the solution. This necessitates an operating temperature of 1000°–1100° C. The electrolyte is molten soda. The cathode, consisting of molten silver, forms 0^{--} ions from the oxygen that is continuously injected. According to the equation $C^{++++} + 2\,0^{--} = CO_2$, the reaction product obtained is carbon dioxide, just as in ordinary combustion. For every carbon atom that is converted, four electrons are given off to the carbon rod and four electrons are withdrawn from the oxygen electrode. These electrons can produce a current in an external circuit. According to this conception, a coal-burning stove is an internally short-circuited fuel cell. The major disadvantage of the fuel cell described above is the high temperature and, consequently, the very short service life of the materials employed. Less severe conditions can be achieved by using gases (hydrogen, in particular) as the fuel. Thus, the Bacon fuel cell (H_2O_2 cell) (Fig. 2) produces current densities of up to about $6\frac{1}{2}$ amp./in.^{2} at a temperature of 240° C. The pressure of the aqueous electrolytes does, however, rise to 1000 lb./in.^{2} and upwards. The ionisation of the gas fed to the cell is effected at diffusion electrodes of nickel. These are porous sintered components which on one side are connected to the gas supply and on the other side are in contact with the electrolyte. The active region is at the boundary of the three phases gas/electrode/electrolyte. To make this boundary as long as possible, all the pores must have the same optimum diameter, as is clarified by Fig. 3 (principle of homoporosity). In order completely to obviate the passage of unutilised gas through the pores, each electrode is provided with a fine-pored cover layer (double-layer electrode). As a result of the high catalytic activity of the electrodes employed, the cell can operate already at room temperature. The H_2O_2 cell designed by Justi and Winsel (Fig. 4), which is known as the dissolved fuel cell, also operates at ordinary temperatures. In this cell the oxygen electrodes contain Raney silver and the hydrogen electrodes contain Raney nickel as the catalyst. Already at temperatures below 100° C (and at atmospheric pressure, too!) this cell attains current density values almost as high as those of the Bacon cell. The fundamental voltage is over 90% of the theoretically attainable voltage of 1.23 volt. The electrodes employed are described as "double-skeleton catalyst electrodes". Because of their great catalytic activity, they are able to dehydrate liquid organic fuels (e.g., methanol). This results in the relatively simple constructional features of the dissolved fuel cell (Fig. 5). The alcohol serving as fuel is mixed with the electrolyte (potassium hydroxide solution).

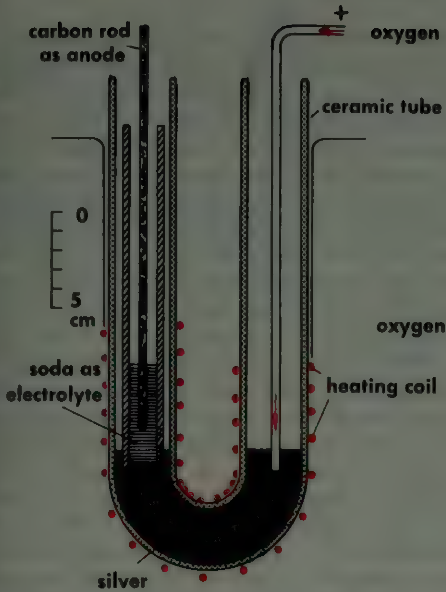


Fig. 1 HIGH-TEMPERATURE CELL ACCORDING TO BAUR AND EHRENBERG

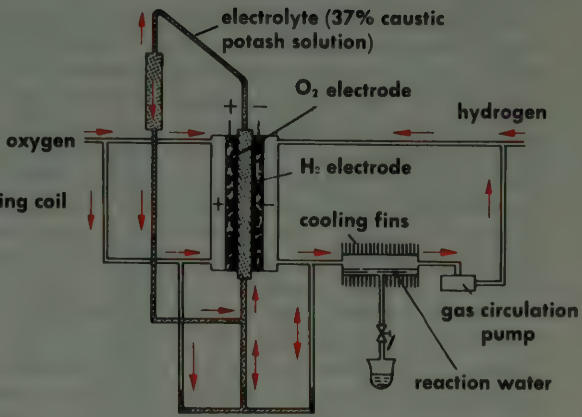


Fig. 2 HIGH-PRESSURE CELL WITH POROUS NICKEL ELECTRODES ACCORDING TO F. T. BACON

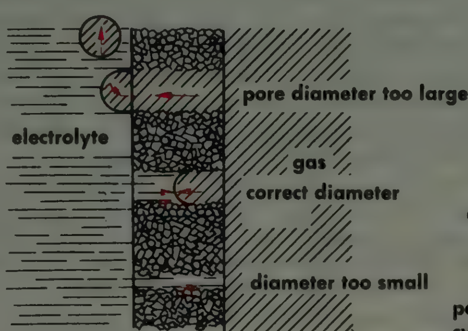


Fig. 3 GAS DIFFUSION ELECTRODE WITH PORES OF DIFFERENT WIDTHS

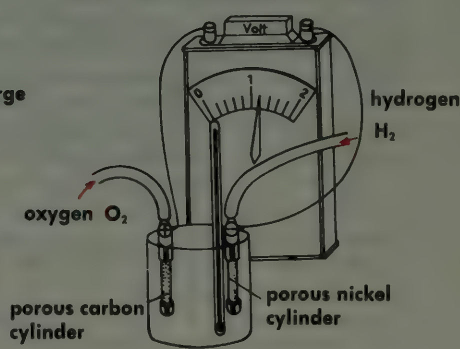


Fig. 4 H_2O_2 CELL ACCORDING TO JUSTI AND WINSEL

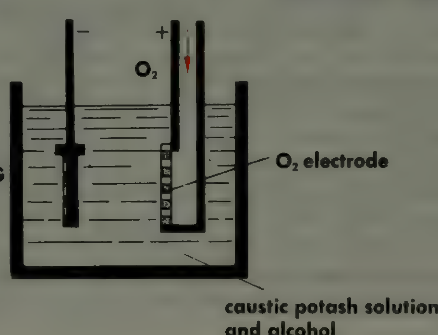


Fig. 5 DISSOLVED FUEL CELL ACCORDING TO JUSTI AND WINSEL
(catalytic dehydrogenation of fuel dissolved in the electrolyte)

The nuclear reactor serves to convert nuclear energy (atomic energy) into thermal energy. The nuclei of atoms consist, broadly speaking, of the elementary particles named protons and neutrons. The protons have a positive electric charge, whereas the neutrons have no charge, i.e., they are electrically neutral. Very powerful forces of attraction act between these particles (collectively referred to as "nucleons") and hold them together in the nucleus. Heavy atomic nuclei are, however, not so stable as light ones, because in the former the repulsive forces exerted by the protons loosen the structure of the nucleus. For this reason it is possible to cause fission of heavy nuclei—such as those of uranium 235—by bombarding them with free neutrons. In Fig. 1 a neutron is shown hitting a nucleus of uranium 235. The latter is set vibrating as a result of this impact, and these vibrations may become so violent that the nucleus is split up into several parts, e.g., into a barium and a krypton nucleus. The "fission products" travel at considerable velocity, collide with matter somewhere in the reactor, and give off their kinetic energy as heat. This is the conversion of nuclear energy into heat. In addition to the fission products and heat formed in the fission of uranium, however, two fresh neutrons are also formed, which in turn can cause the fission of more uranium atoms. In this way the *chain reaction* is initiated; this is illustrated in Fig. 2: the neutron coming from the left strikes the U 235 nucleus and briefly forms the intermediate product U 236, which disintegrates spontaneously into strontium and xenon. Three neutrons are released in this fission process. In order to be able to utilise these neutrons, which are emitted from the parent nucleus at high velocity, for further fissile processes, they have to be slowed down ("moderated"). Low-velocity neutrons are much better suited to split atoms than high-velocity neutrons are. The slower neutrons can interact with the uranium nucleus for a greater length of time, whereas faster neutrons are in the vicinity of the nucleus for too short a time to initiate the fission process. The velocity of the neutrons is moderated by causing them to collide with light atoms, large numbers of which must be incorporated in the reactor for this purpose. Materials consisting of such light atoms are, for example, graphite and water. The neutrons which have been slowed down in this way will then cause fission of further U 235 nuclei. Each fission process gives birth to fresh electrons, so that the chain reaction is self-sustaining and the reactor is consequently kept in operation.

Fig. 3 is a schematic illustration of a nuclear reactor—in this case a water-moderated reactor. The actual reactor core is shown on the left. The uranium is installed in the form of metallic rods in a vessel filled with water. The fission takes place within the uranium. The neutrons which are released in the process are diffused through the uranium and into the surrounding water, where they collide with the light hydrogen and oxygen atoms and are moderated, i.e., they lose velocity.

These slowed-down neutrons re-enter the uranium rods with a certain probability and they there cause fresh fission reactions to take place. The fission products formed as a result of these give off their energy to the uranium, which in turn transmits it to the water. The latter is circulated by a pump through a heat exchanger, in which it transfers the heat to a secondary thermal circuit. To ensure that the reactor will not stop functioning nor become excessively overheated, the rate of neutron formation inside it has to be controlled with considerable accuracy. This is done by means of the control rods, which consist of a neutron-absorbing material and which are inserted into the reactor core to an accurately variable depth. This depth of penetration must be just enough to ensure that, on an average, per fission only one neutron remains available to produce another fission reaction. Since the fission products are highly radioactive, the reactor must be enclosed in a thick casing of concrete called the "shield". In the so-called boiling water reactor (Fig. 4) the water is allowed to evaporate inside the reactor itself, the steam being utilised in the primary circuit; in principle, no heat exchanger is needed.

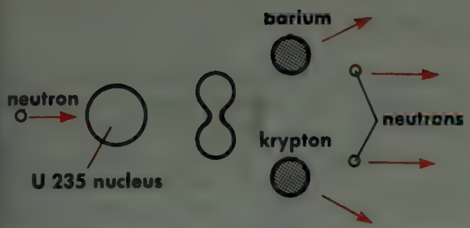


Fig. 1 FISSION OF A URANIUM 235 NUCLEUS BY A NEUTRON

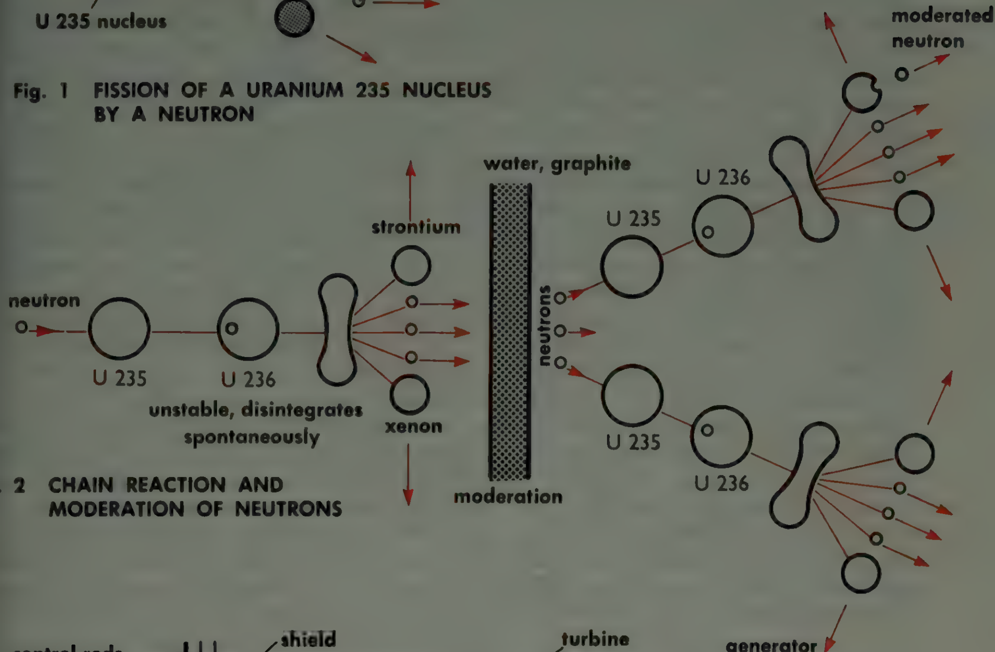


Fig. 2 CHAIN REACTION AND MODERATION OF NEUTRONS

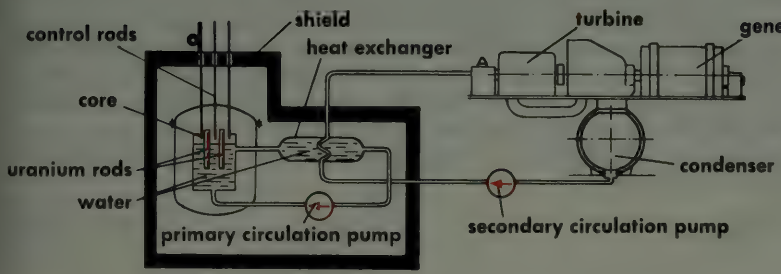


Fig. 3 WATER-MODERATED REACTOR (*schematic*)

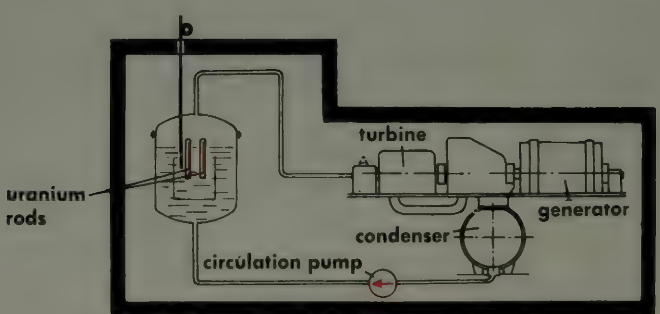


Fig. 4 BOILING WATER REACTOR (*schematic*)

Matter is composed of neutral atoms. The electrical neutrality of the atoms is due to the fact that the positive charge of the nucleus of the atom is compensated by the negative charge of the electrons that surround it. The outermost electron may either be only loosely connected to the rest of the atom (Fig. 1a) or be more firmly embedded in it (Fig. 1b). Atoms of the first type tend to part with electrons to adjacent atoms, while those of the second type tend to tear electrons away from adjacent atoms. It is because of this phenomenon that, for example, glass becomes positively charged when it is rubbed with a silk cloth (Fig. 2a), whereas ebonite acquires a negative charge on being rubbed with a woollen cloth (Fig. 2b). If two small balls, suspended pendulum wise, are given electric charges of the same sign (e.g., by touching each ball with a glass or ebonite rod charged by friction in the manner described), they will be pushed some distance apart by the mutually repellent force exerted by the two similarly charged balls, as in Figs. 3a and 3b. On the other hand two oppositely charged balls will attract each other, as in Fig. 3c, and when they come into contact, their charges will neutralise each other. A positive charge means that there is a deficiency of electrons; a negative charge means that there is a surplus of electrons in relation to the neutral condition of the atoms. Electrons are the elementary particles of electricity. Each electron has a charge $e = 1.602 \times 10^{-19}$ Coulomb, a rest mass $m_0 = 0.9108 \times 10^{-27}$ grammes, and a radius of 2.82×10^{-13} cm.

Since similar charges repel each other, the electrons so arrange themselves at the surface of an electrical conductor (e.g., a metal ball) that the space inside it contains no charge and is thus entirely free from electrical forces (Figs. 4a and 4b). When a charged conductor is brought near an uncharged (neutral) one, a separation of charges is induced on the latter (Fig. 5a). If the first conductor has a positive charge, then the initially neutral conductor will become negatively charged on the side facing the positive conductor, while the other side will acquire a positive charge of the same magnitude. This positive charge can be discharged to the earth (Fig. 5b), and the second, initially neutral conductor will then remain negatively charged (Fig. 5c). This method of charging is known as electrostatic induction. The repulsion of similar charges may produce repellent forces of considerable magnitude at pointed extremities, where the electrons become concentrated and are actually discharged from the conductor (Fig. 6), so that they can, as it were, be sprayed on to a neutral conductor, which in turn will acquire a charge of the same sign (in this case negative) as that of the first conductor.

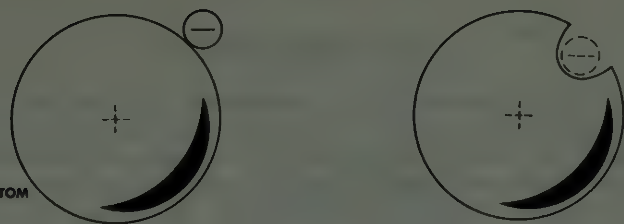


Fig. 1
BOND OF ELECTRON TO ATOM

Fig. 2 FRICTIONAL ELECTRICITY

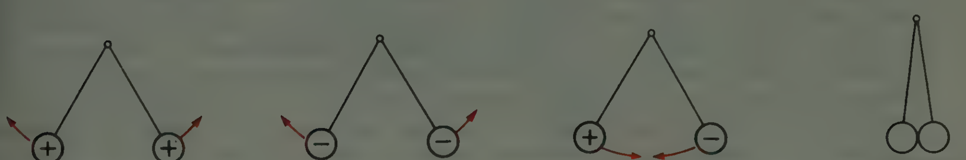
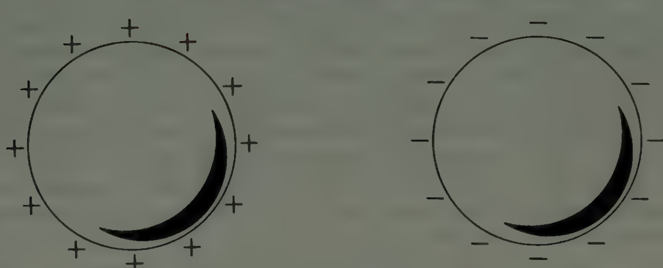


Fig. 3 FORCES EXERTED UPON ELECTRICALLY CHARGED BALLS

ig. 4
ISTRIBUTION OF CHARGE
ON THE SURFACE



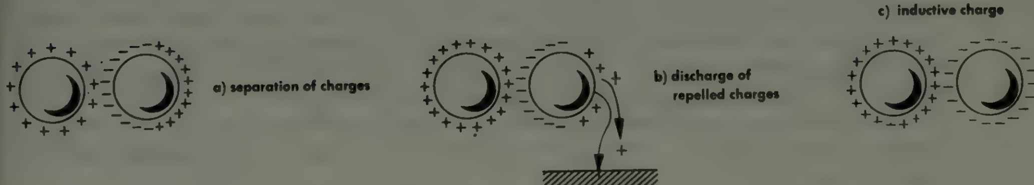
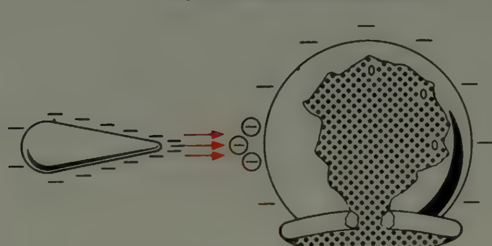


Fig. 5 ELECTROSTATIC INDUCTION

Fig. 6 POINT EFFECT



The Van de Graaff generator (Fig. 1) makes use of the possibility of spraying an electric charge from sharp points of a conductor. The charge is thus applied to a belt conveyor made of an insulating material, which conveys the charge into the interior of a spherical conductor of large radius, where the charge is collected by a "comb" of pointed electrodes. More and more electricity is collected in this way from the travelling belt and is accumulated at the surface of the large sphere, which thus acquires a very high charge. The power output from this electrostatic generator is not very large, for the charge accumulated by this method cannot sustain a current of any significant magnitude. On the other hand, very high voltages can be obtained (of the order of some millions of volts). The voltage can be further increased by installing the generator in an enclosed space in which the air pressure is increased above the normal atmospheric pressure, so that the spark-over voltage to earthed components is increased. The amount of electric charge that can be stored up in a body is called the *capacity* of that body. A *condenser* (or capacitor) is a device specifically intended to store up an electrical charge. Its capacity is determined mainly by the action of electrostatic induction. It consists essentially of two conducting surfaces (plates) which are insulated from each other (Fig. 2a). In the case of a variable condenser the area (F) of these surfaces and/or their distance apart (d_1, d_2) can be varied. Obviously, the quantity of electricity that can be stored up by induction will be greater according as F is larger and the gap d between the condenser plates is smaller. The capacity of a plate condenser is therefore proportional to F and inversely proportional to d. High-capacity condensers are composed of "plates" consisting of rolled-up thin metal foils separated by sheets of paper as the insulating medium. The capacity of a condenser may be compared with the cubic capacity of a tank, which depends on the area of the bottom and on the height (Fig. 2b). If a small ball pendulum is attached to a conductor and the latter is charged with electricity, the pendulum will acquire a deflection (due to electrical repulsion) which is proportional to the magnitude of the charge (Fig. 3a). There is an analogy with the pressure of water in a tank, which pressure can be measured by means of a mercury manometer (or pressure gauge) (Fig. 3b). The pressure of the water corresponds to the electric potential or voltage (the unit of measurement being the volt). The voltage (U) is associated with the electric charge Q (measured in coulombs) and the capacity C of a conductor (measured in farads) by the following relation: $U = Q/C$. In the space which surrounds an electrically charged body an electric potential occurs which is proportional to the charge Q and inversely proportional to the distance r from the centre of the body ($U \approx Q/r$). The electrical condition produced in a space by the presence of electrically charged bodies is called an *electric field* (Fig. 4). Points which all have the same potential (voltage) are located on equipotential surfaces. Forces always act in the direction of the potential gradient. The electric force which is exerted upon a charge of unit magnitude in an electric field is called the field strength (or field intensity). It is always directed perpendicularly to the equipotential surfaces. The lines of force in an electric field represent the direction of the force at any point on their length. The properties of an electric field can be described in terms of the equipotential surfaces and lines of force. The lines of force are conceived as emerging from positive charges and disappearing into negative charges.

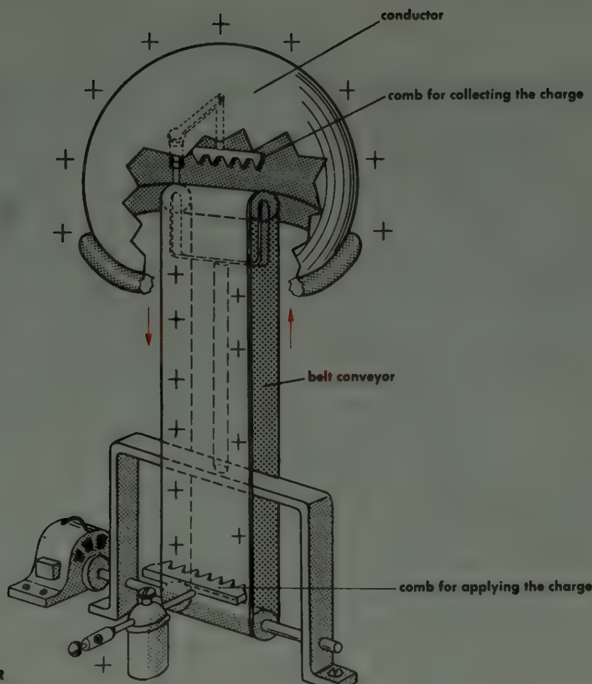


Fig. 1 VAN DE GRAAFF GENERATOR

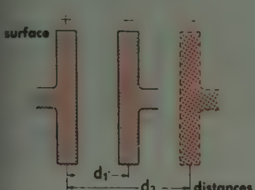


Fig. 2a ELECTRIC CAPACITY

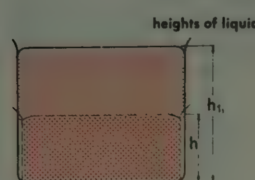


Fig. 2b HYDRAULIC CAPACITY

comparison between electrostatic and hydrostatic phenomena

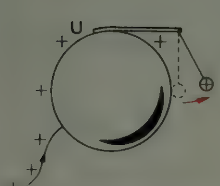


Fig. 3a ELECTRIC VOLTAGE U

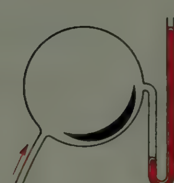


Fig. 3b HYDROSTATIC PRESSURE

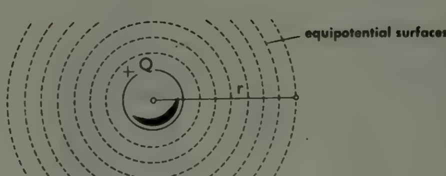


Fig. 4 ELECTRIC FIELD

If two electrostatically charged conductors differing in potential ($U_1 > U_2$) are interconnected by means of a metal wire, charge equalisation in the form of an electric current will occur until the two conductors have acquired the same potential (Fig. 1a). This process is comparable to the flow which occurs when a tank containing a liquid is connected by means of a pipe to another tank, situated at lower level, so that there is what is known as a pressure gradient between the two tanks (Fig. 1b). To obtain a sustained flow of electricity – i.e., an electric current – it is necessary to maintain the potential difference between the two conductors. In the case of the Van de Graaff generator (see page 72) this is achieved by the continuous input of electric charge by the belt conveyor (Fig. 2a). In the hydraulic analogy of the two tanks the same effect would be produced by a pump (or, to make the analogy even closer, by a bucket elevator which raises the liquid from the lower level to the higher level) (Fig. 2b). However, the flow or electricity differs from the flow of a liquid in that it produces a magnetic field around it. The relation between the direction of current and the polarity of the magnetic field is conveniently expressed by the "corkscrew rule": A current following the direction of twist of a corkscrew (clockwise direction) produces magnetic lines of force in the direction of its thrust, i.e., the south pole faces the observer. Conversely, if a current flows through a conductor in the direction of thrust of the corkscrew, the direction of rotation ("screwing-in" direction) corresponds to the direction of the magnetic lines of force surrounding that conductor. The presence of this magnetic field can be demonstrated, for example, by the deflection of a compass needle (Fig. 3a). The effect of an electric current flowing through a coil of wire is to produce a magnetic field similar to that of a bar magnet (Fig. 3b). The magnetic effect of a coil can be intensified by causing the magnetic flux (i.e., the lines of force conceived as a "flow" from one pole of the magnetic to the other) to pass through a core consisting of a material whose magnetic resistance is lower than that of air. The most suitable substance for the purpose is iron. A coil provided with an iron core is called an electromagnet (Fig. 4a). Frequently two electromagnets are joined together by a yoke, i.e., an iron bar forming a magnetic path (Fig. 4b). In an arrangement of this kind the two coils are wound in opposite directions so that they have a free north pole and a free south pole respectively. The magnetic lines of force are conceived as flowing from the north pole to the south pole. Magnetism is always associated with dipoles, i.e., there are always two poles; there are no free positive and negative magnetic charges similar to electrostatic charges.

(Continued)

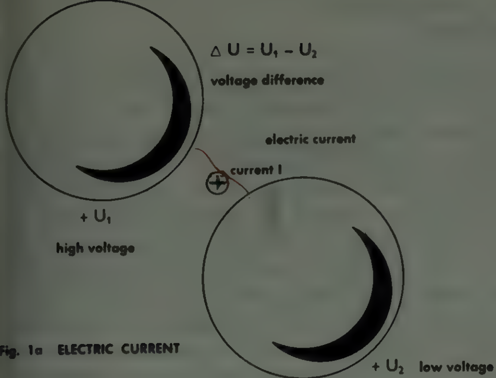


Fig. 1a ELECTRIC CURRENT

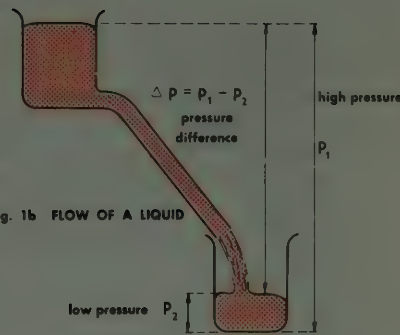
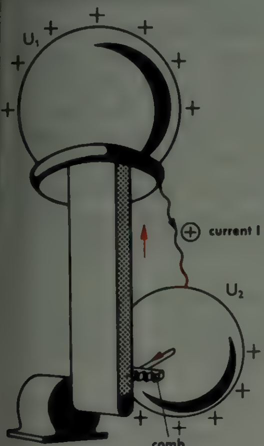


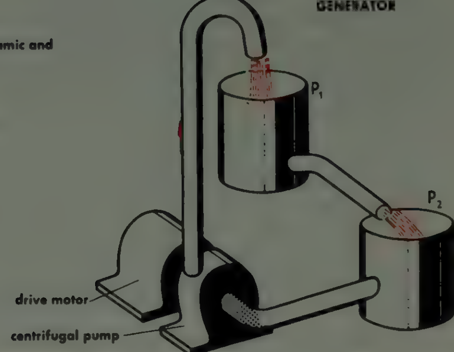
Fig. 1b FLOW OF A LIQUID

Fig. 2a ELECTROSTATIC GENERATOR



comparison between electrodynamic and hydrodynamic phenomena

Fig. 2b HYDRAULIC GENERATOR



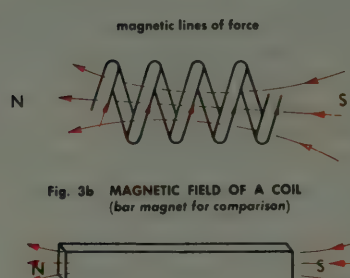


Fig. 3b MAGNETIC FIELD OF A COIL
(bar magnet for comparison)

Fig. 3a MAGNETIC FIELD OF A CONDUCTOR

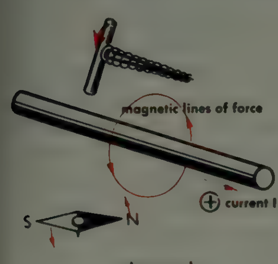


Fig. 4a ELECTROMAGNET

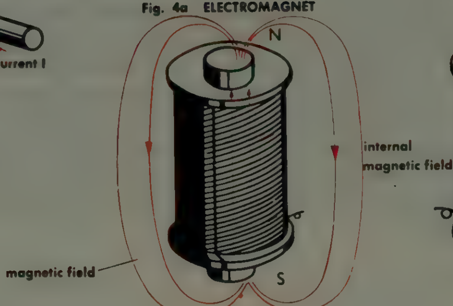
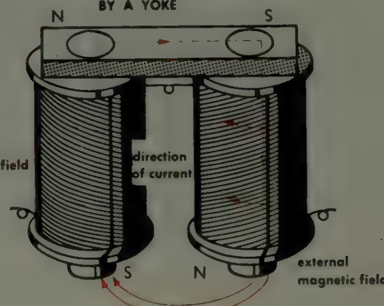


Fig. 4b TWO ELECTROMAGNETS JOINED BY A YOKE



The phenomena associated with the flow of electricity in metallic conductors can be explained with the aid of the interaction of the elementary electric charge of the electron with the atoms of the metals. For convenience, the electrons are assumed to have a spherical shape. A metal very widely employed for the conduction of electricity is copper. It has a crystalline structure (Fig. 1). The nucleus of the copper atom contains 29 positive elementary charges, which are neutralised by 29 negatively charged electrons. The 29th (outermost) electron is only very loosely connected to the atomic nucleus. Even at room temperature the thermal energy is great enough to enable the copper atoms to perform vibrations about their position of rest in the crystal lattice. As a result, these loosely connected electrons are, as it were, shaken off and thus become available as free carriers of negative electric charge for the conduction of electricity. These electrons are "quasi-free", i.e., they are repeatedly captured and released again. In the crystal lattice they behave rather like a gas in a container; for this reason the term "electron gas" is sometimes employed (Fig. 2). When a potential difference is applied between the ends of a conductor, electrons go from the negative to the positive pole (Fig. 3). The flow of electrons thus moves in the direction opposite to that of the current as conventionally defined. The behaviour of an electron stream in a magnetic field is shown in Fig. 4, where a single individual electron (e.g., emitted by an incandescent wire in a vacuum) traverses a constant magnetic field. The electron itself is surrounded by its own magnetic field, which is superimposed upon the main field. Under the conditions considered, the magnetic field strength above the electron path is thereby intensified, and the field strength below it is reduced. The resulting field strength gradient causes the electron to move in a curved path. In the interior of a metal conductor this gives rise to a difference in potential, or voltage, between the upper and the lower face of the conductor (Fig. 5). This phenomenon is called the Hall effect, after its discoverer.

The electron theory of metallic conductivity also enables *induction* to be explained in terms that can be visualised. By induction is understood the occurrence of electric potential differences and currents as a result of mechanical movements of conductors in a magnetic field.

As shown in Fig. 6, a piece of metal can be conceived as a kind of container which is filled with "electron gas" and which, for the purpose of the present explanation, is assumed to be moving in a constant magnetic field. As a result of this motion the electrons in the metal will tend to follow downward-curving paths and thus concentrate at the lower end, so that this end will acquire a negative potential in relation to the upper end of the piece of metal. If the two ends are interconnected by a wire which extends far outside the magnetic field, an induction current will flow through this wire so long as the piece of metal continues to move through the magnetic field. In general, the current will flow so long as the amount of magnetic flux intersecting the plane of a circuit (which in its simplest form is merely a loop of wire) varies. The variation of the flux can be obtained either by varying the area enclosed within the circuit (this is the principle applied in the electric generator) or by varying the magnetic field strength and thus causing changes in the density of the flux (as in the transformer: see page 110). The operating principle of the *generator* is illustrated in Fig. 7a: The amount of magnetic flux passing through the plane of the rotating loop of wires varies periodically from zero to its maximum value, according as the plane of the loop is parallel or perpendicular to the magnetic lines of force. An electric current which varies periodically in direction (*alternating current*) is thereby induced in the loop. This current is collected by means of contact rings (known as slip-rings) and flows through the external circuit (Fig. 7b). The collector may also be in the form of a so-called commutator consisting of a number of separate segments (Fig. 7c, showing only two segments), whereby a (pulsating) *direct current* is obtained. See also pages 78 and 82.

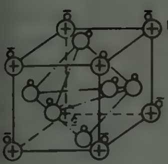


Fig. 1 CRYSTALLINE STRUCTURE OF COPPER

Fig. 2 ELECTRON GAS

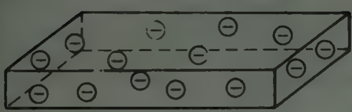


Fig. 3 FLOW OF ELECTRONS

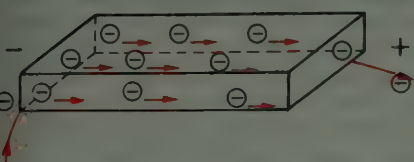


Fig. 4 ELECTRON IN A MAGNETIC FIELD

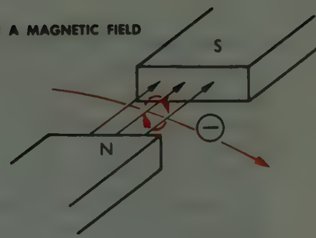


Fig. 5 HALL EFFECT

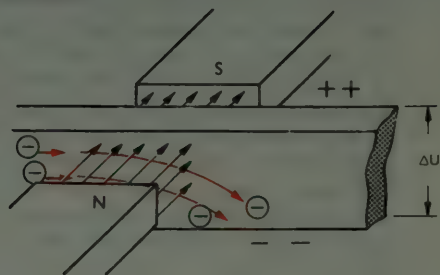
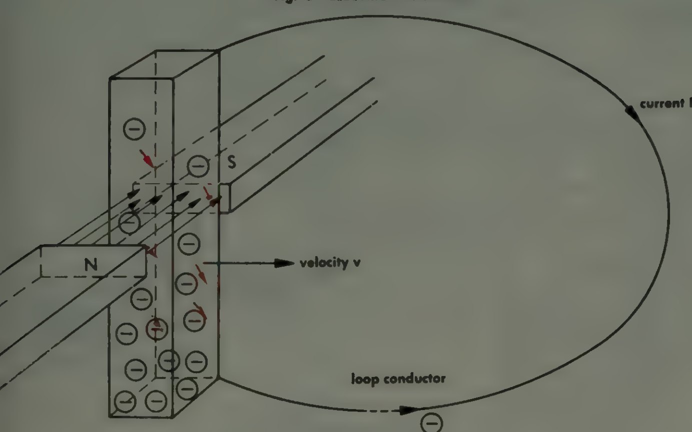


Fig. 6 ELECTRIC INDUCTION



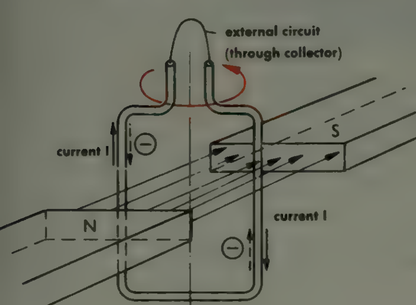


Fig. 7a ELECTRIC GENERATOR
(dynamo)

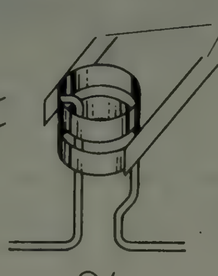


Fig. 7b SLIP-RING COLLECTOR
(alternating current)

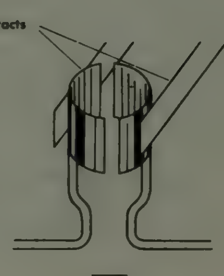


Fig. 7c COMMUTATOR
(direct current)

ELECTRIC GENERATOR (DYNAMO)

Generators are machines used for the large-scale production of electrical energy. Their operation is based on the principle of electrical induction (see also page 76), whereby a periodic flow of electricity is produced in a loop-type conductor as a result of the periodic variation of the flux of the magnetic lines of force passing through this loop. To do this, we can either cause the loop to rotate in a constant magnetic field or, alternatively, the loop can be kept stationary and the magnetic field rotated. In the former arrangement the "loop" is formed by the armature windings on the rotor which revolves between the fixed magnetic poles of the stator. In the latter arrangement the armature is stationary, and the magnetic poles (on a so-called "magnet wheel") revolve instead; the stator consists of an iron ring with induction coils mounted on the inside; the magnetic poles on the rotor move past the ends of these coils at a very short distance from them (Figs. 1 and 3). In this case the current produced by the generator is taken direct from the stator, without the aid of special current collectors (brushes). For this reason this form of construction is particularly suitable for the generation of high-voltage alternating current. The sparking that occurs at high voltages (around 20,000 volts) in large generators would destroy the brushes. The relatively low output of direct current needed for producing the rotating magnetic field is fed to the rotor by means of slip-rings and carbon or copper-mesh brushes (Fig. 3). The successive coils in Fig. 1 are wound in alternate directions, which ensures that the generated current always flows in the same direction. High-duty generators are usually coupled directly—on the same shaft—to steam or water turbines. Usually, a small direct-current dynamo for producing the magnetic field is also mounted on the driving shaft (Fig. 2). In the older type of power station with reciprocating steam engines, the rotor of the generator is generally constructed as a flywheel with the magnetic pole windings round its rim. Fig. 3 shows a smaller generator which likewise operates on the principle described above (rotating magnetic field, stationary armature winding). In this case the magnet wheel is in the form of a two-part T-rotor.

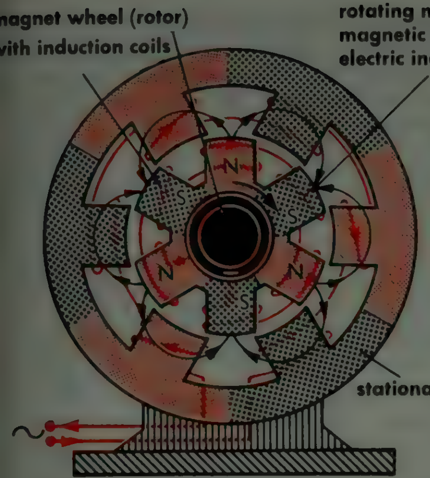


Fig. 1 ALTERNATING-CURRENT GENERATOR
(*internal pole machine, schematic*)

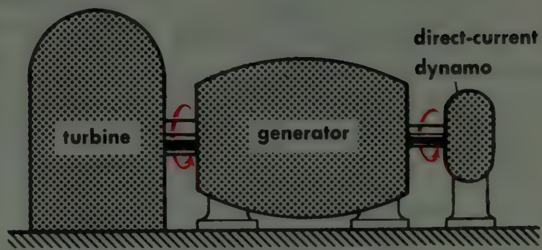


Fig. 2 GENERATING SET

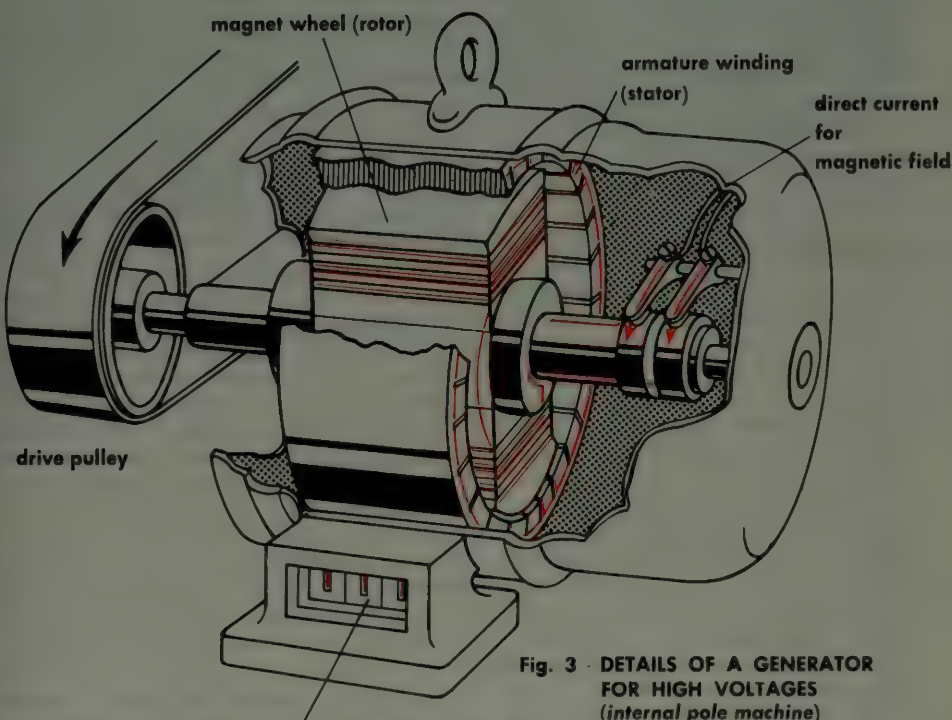


Fig. 3 DETAILS OF A GENERATOR
FOR HIGH VOLTAGES
(*internal pole machine*)

If two different metals are immersed in an aqueous solution—i.e., a solution of a substance in water—which can conduct electricity (e.g., dilute sulphuric acid), they will have a different tendency to dissolve in the liquid. The two metals thereby acquire electric charges of different sign. As a result of this chemical process, a difference in voltage occurs because one of the metals appears positive or negative in relation to the other. In this respect the various metals or other conductors can be arranged in a so-called electrochemical or electromotive series in which each metal has a higher potential—i.e., appears electrically positive—in relation to the next metal: carbon, gold, silver, copper, tin, lead, iron, zinc.

A combination of two metals (the electrodes) in an aqueous solution for producing electrical energy from chemical energy is sometimes called a galvanic cell. One such cell is the voltaic cell (Fig. 1) in which the two metals are copper (+) and zinc (−) and the solution in which they are immersed is dilute sulphuric acid. If an incandescent lamp (for example) is connected to the two poles of the cell, an electric current will flow. The lamp and connecting wires form the external part of the electric circuit, which is completed in the interior of the cell by the conducting liquid (which is called the electrolyte). The flow of current inside the cell causes polarisation, i.e., bubbles of gas are deposited on the electrodes, with the result that the internal resistance is increased and the flow of current is, after a time, greatly reduced. In particular, the hydrogen gas that is evolved at the anode (the positive pole) must be removed if the cell is to go on functioning properly. In the Leclanché cell (Fig. 2), the undesirable polarisation is prevented by a mixture of manganese dioxide and graphite as depolarisers which, in a simple form of this cell, are contained in a linen bag which encloses the carbon anode. Cells of this kind are nowadays sometimes used for supplying the current to work an electric bell. It consists of a carbon anode with the depolariser, a zinc cathode (negative pole), and an ammonium chloride solution as the electrolyte. The potential difference between the poles of the Leclanché cell is 1.3 volts. The size of the cell does not affect this voltage, but it does determine the intensity of the current (amperage) it can give. The chemical energy that is converted into electrical energy when the cell is in operation is obtained from the zinc electrode, which dissolves and is thus consumed. This process is irreversible, and the zinc must therefore be renewed from time to time. The depolariser is also consumed and requires occasional renewal. Cells in which the electrodes are consumed are called primary cells. On the other hand, a secondary cell can be restored to its original state by charging, i.e., passing an electric current through it, so that the electrodes are regenerated. Cells of this latter type, usually grouped in batteries of two or more, are called accumulators (storage batteries). A commonly used type of accumulator has electrodes consisting of lead plates; the electrolyte is dilute sulphuric acid (Figs. 3 and 4). A layer of lead sulphate is formed on these plates. When the accumulator is charged, the layer on the anode plate is transformed into brown lead dioxide, while the cathode is reduced to grey lead. Electrical energy (the charging current) is converted into chemical energy. In the charged accumulator, one electrode thus consists of lead and the other of lead dioxide; these electrodes, together with the electrolyte, then function as a voltaic cell. The stored (accumulated) chemical energy is converted back into electrical energy (discharging). The accumulator is widely used for technical purposes. Besides the familiar lead accumulator there are several other kinds, including the nickel–iron (Ni–Fe) accumulator, which has potassium hydroxide as electrolyte. The lead accumulator produces a potential difference of about 2 volts; for the nickel–iron accumulator it is 1.36 volts. Secondary as well as primary cells are also manufactured as so-called "dry" cells, in which the electrolyte is in the form of a paste instead of a liquid. Higher voltages are obtained by connecting two or more cells in series (Fig. 5), higher current intensities by connecting them in parallel (Fig. 6). All cells produce direct current, i.e., an electric current which flows constantly in one direction (as opposed to alternating current: see page 76). The term "battery" denotes a group of two or more primary or secondary cells, connected in series or in parallel.

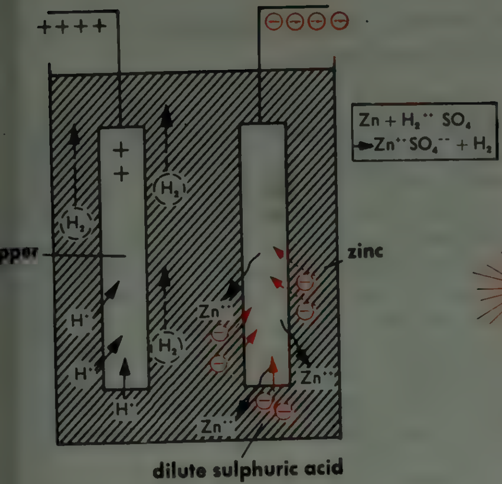


Fig. 1 VOLTAIC CELL

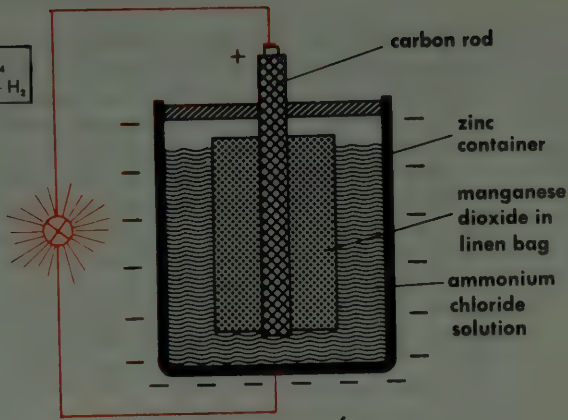


Fig. 2 LECLANCHÉ CELL

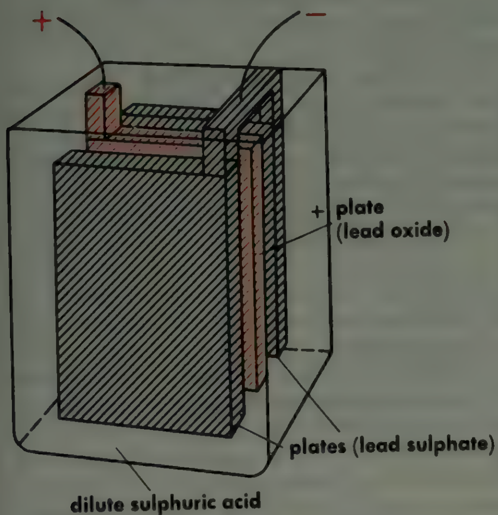


Fig. 3 ACCUMULATOR (*principle*)

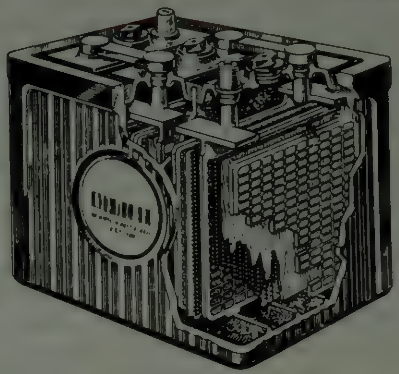


Fig. 4 ACCUMULATOR (*automobile battery*)

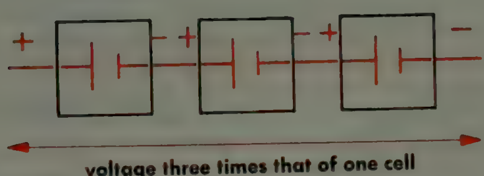


Fig. 5 CELLS CONNECTED IN SERIES

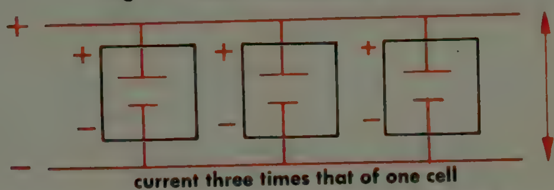


Fig. 6 CELLS CONNECTED IN PARALLEL

ALTERNATING CURRENT, THREE-PHASE CURRENT, ELECTROMAGNETIC WAVES

When electrical energy is produced by induction in a generator, a primary alternating voltage is produced which causes an alternating current to flow in an external circuit (Fig. 1a; see page 78). The presence of capacity and self-induction in the circuit may cause phase displacements between the voltage and the current (Fig. 1b). The product of voltage and current is the power. In a case where the voltage and the current have the same algebraic sign, the power is positive (Fig. 1a). If a phase displacement between current and voltage occurs, they are liable to be of different sign, in which case the power becomes negative. Because of these negative values, the active power (true power) of alternating current which is not in phase (Fig. 1b) is smaller than that of current which is in phase (Fig. 1a). If the current is displaced a quarter period in relation to the voltage, then there will be no active power at all; in that case only the so-called idle current (wattless current) flows through the circuit. The reduction in power that occurs as a result of phase displacement is expressed by the power factor $\cos \phi$ (electric power $N = U. I. \cos \phi$).

If three alternating currents have phase displacements of 120° in relation to one another, then at any particular instant the sum of their currents or of their voltages will be zero (Fig. 2a). Instead of six, only three conductors are required for the transmission of these three currents if either "star" connection (Fig. 2b) or "delta" connection (Fig. 2c) is employed for the conductors.

If three electromagnets arranged as shown in Fig. 2d are energised by three such alternating currents (which together are referred to as "three-phase alternating current"), the maximum values of the current in the three electromagnet coils will occur successively, with a phase displacement of 120° in relation to one another. This has the effect of producing a rotating magnetic field. A so-called "squirrel-cage" rotor, for instance, can be made to rotate in such a magnetic field, so that this principle can very suitably be applied to the construction of three-phase electric motors.

Alternating currents are characterised by their frequency, i.e., the number of cycles (or double waves) per second. The term "period" signifies the time occupied by one complete cycle. The unit of frequency is the hertz, which is equal to one cycle per second. Low-frequency alternating current is a term applicable to frequencies up to 20,000 hertz.[1]) Ordinary alternating current from the mains has a frequency of 50[2] hertz. The electric railways in certain countries operate at a frequency of 16⅔ hertz. High-frequency alternating currents up to several gigahertz (1 gigahertz = 1 milliard[3] hertz)—are used in communication engineering and more particularly in radio communication. This is because an alternating current produces an electromagnetic field which varies with the frequency of the current and which is propagated through space with the velocity of light, i.e., 180,000 miles per second. With low-frequency alternating current the energy that is transmitted into space has sufficient time to return to the electrical conductor each time the alternating current changes its direction of flow. With very high frequencies, on the other hand, the change of direction takes place before the whole of the energy has had time to return. A proportion of the energy is therefore, as it were, cut off and is transmitted into space as *electromagnetic radiation.* Because of the periodic character of this radiation, it is referred to as an "electromagnetic vibration", and because of its similarity with the propagation of a wave in water, it is also called an "electromagnetic wave".

(Continued)

1. The usual subdivision of frequency ranges is: low frequency up to 20,000 hertz; medium frequency from 20,000 to 30,000 hertz; high frequency from 30,000 to 3000 megahertz (= 3000 million hertz); beyond this is the range known as hyperfrequency.
2. 60 in U.S.A.
3. One billion in U.S.A.

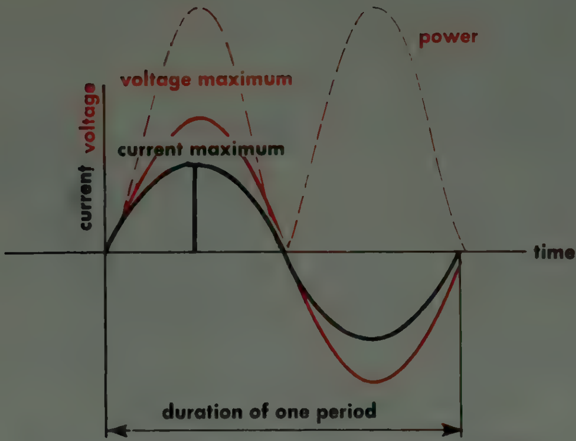


Fig. 1a TIME GRAPH OF AN ALTERNATING CURRENT

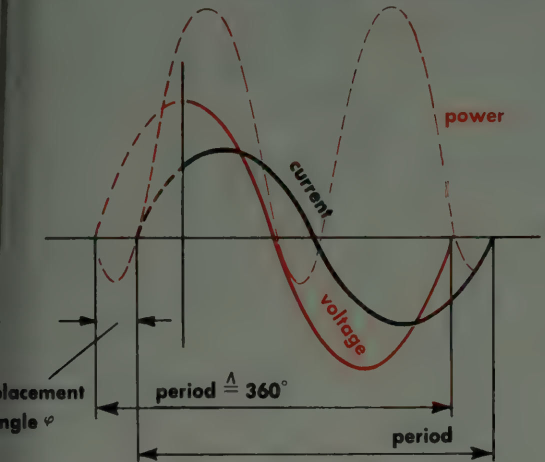


Fig. 1b PHASE DISPLACEMENT BETWEEN CURRENT AND VOLTAGE

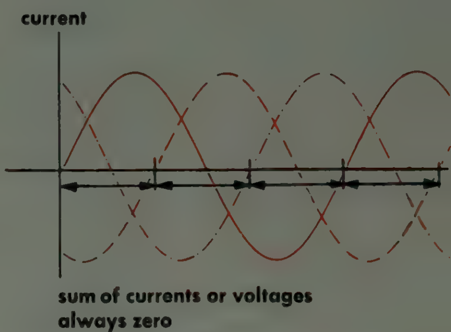


Fig. 2a INTERLINKING OF THREE ALTERNATING CURRENTS

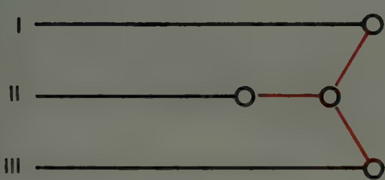


Fig. 2b STAR CONNECTION

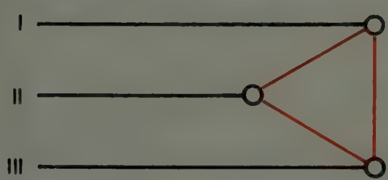
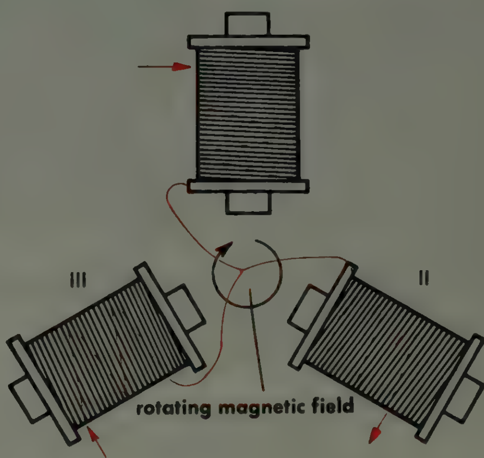
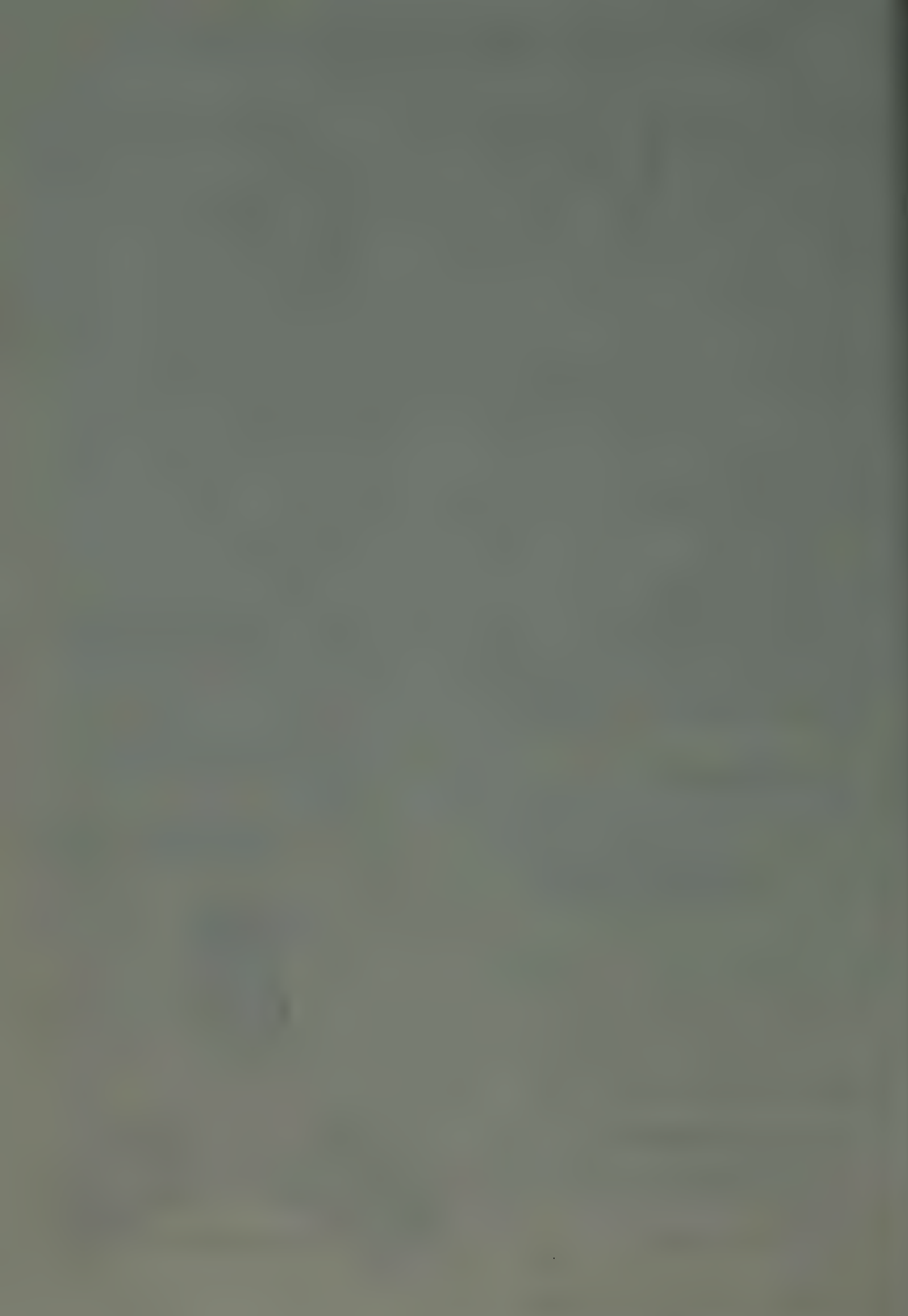
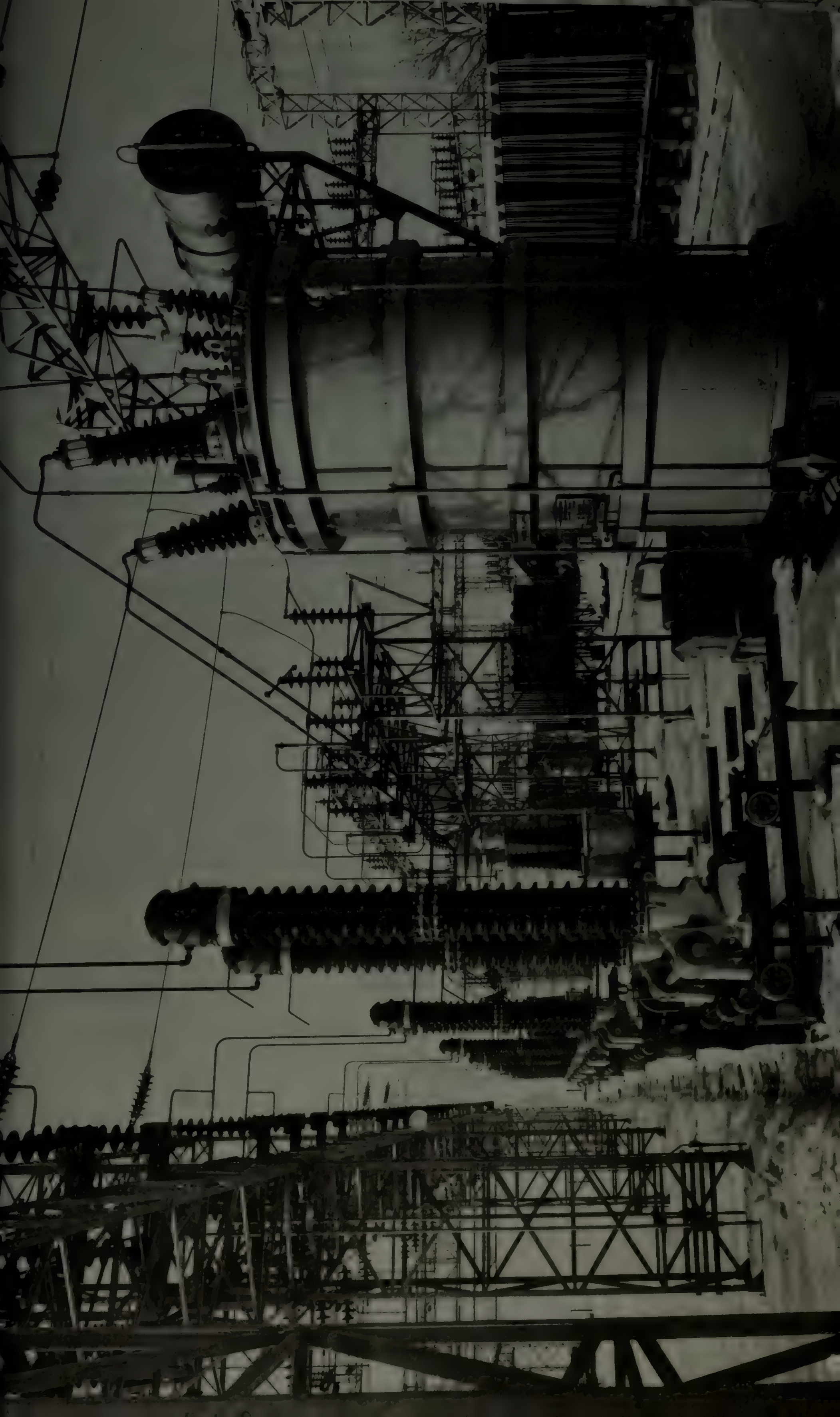


Fig. 2c DELTA CONNECTION





High-voltage accumulator for Swiss Railways
Photo CFF



The generation of electromagnetic waves is therefore always associated with the existence of a high-frequency alternating current whose energy is radiated into space. This radiation process, which depends on the frequency of the alternating current that produces it, can best be demonstrated with reference to the so-called Hertz dipole (Fig. 3a). This device consists of a conductor whose length is related to the length of the period—more commonly known as the *wavelength*—of the electromagnetic wave produced. The wavelength is found as the velocity of propagation (= distance travelled per second) divided by the frequency (= number of periods or cycles per second). If a high-frequency alternating current is passed through this conductor (or is produced in the latter by induction), a high-frequency alternating electromagnetic field will be formed around it, as shown in Fig. 3a. The emission of electromagnetic waves at the critical distance r_k is illustrated in Fig. 3b. If one end of a dipole is earthed, we have the simplest form of an *aerial* (antenna). A circuit whose electrical and magnetic properties (capacity and self-induction respectively) are tuned to the frequency of the alternating current—i.e., is in resonance with it—is called an *oscillating circuit*. The latter comprises a condenser and a self-induction coil. Once it is charged, the condenser discharges itself through the coil, so that electrical energy that was stored up in the condenser is converted into magnetic energy. After the discharge of the condenser has taken place, the magnetic field breaks down and induces a current in the coil. This current gives the condenser a charge of opposite sign to its original charge. If no losses occurred, the charge would go on oscillating—changing its sign—indefinitely. The frequency of this oscillation, and therefore the frequency of the alternating current that is produced, is higher in proportion as the capacity C and the self-induction L are lower (Fig. 4a). The oscillation period T is given by the formula $T = 2\pi\sqrt{LC}$.

If the plates of the condenser are separated, as in Fig. 4b, they can be conceived as the "aerial" and "earth" respectively. This principle of the aerial (antenna) is further illustrated in Figs. 4c and 4d. In these cases the electric field of the condenser is utilised for the transmission and reception of the electromagnetic energy in the form of waves. A coil wound round an open frame (directive aerial) (Fig. 4e) can be used for receiving the magnetic portion of the field. The frame is rotated to the position where maximum variation of the magnetic flux is obtained.

An important piece of equipment for producing hyper-frequency oscillations (see page 92 *et seq.*) is the cavity resonator. Fig. 4f shows how such a resonator can be conceived as being evolved from an oscillating circuit of minimum capacity and self-induction. The lateral wall of the "box" corresponds to the self-induction and the end surfaces correspond to the capacity.

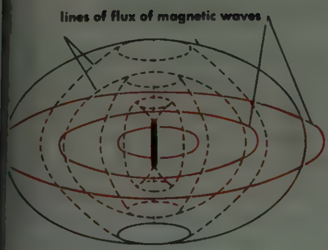


Fig. 3a SPHERICAL WAVE SYSTEM AROUND A HERTZ DIPOLE

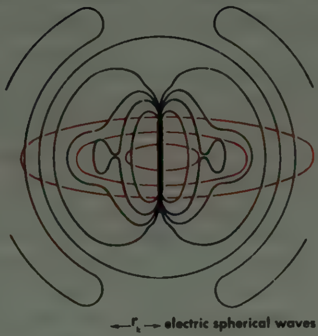


Fig. 3b EMISSION OF WAVES

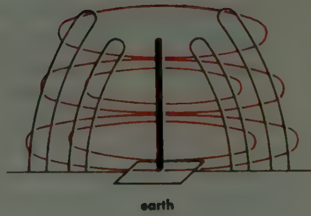


Fig. 3c DIPOLE EARTHED ON ONE SIDE (*antenna*)

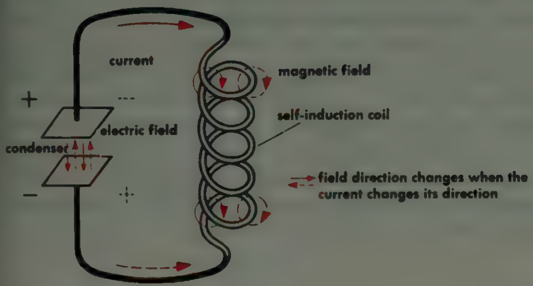


Fig. 4a CLOSED OSCILLATING CIRCUIT

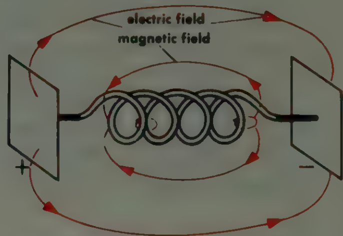


Fig. 4b OPEN OSCILLATING CIRCUIT

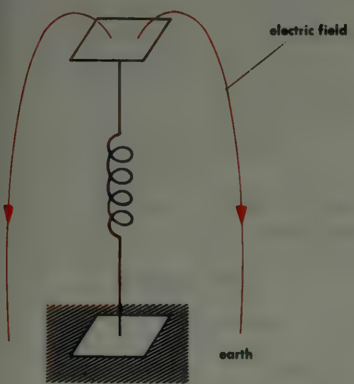


Fig. 4c ANTENNA

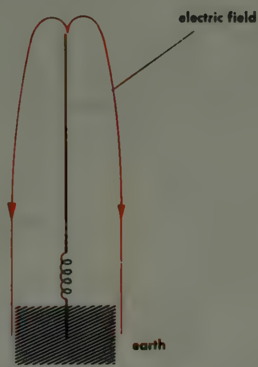


Fig. 4d ANTENNA

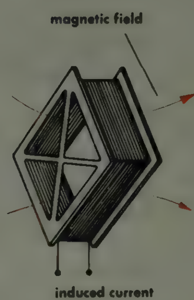


Fig. 4e DIRECTIVE ANTENNA

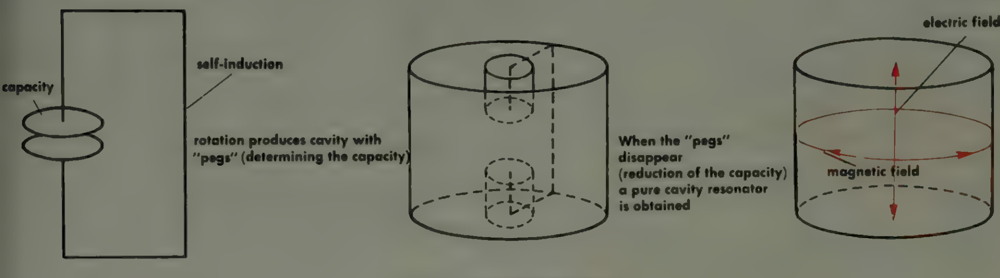


Fig. 4f CAVITY RESONATOR

Oscillating circuits are used, in the form of aerials, as transmitting and receiving devices for electromagnetic energy in radio transmitters and receivers. As the emission of electromagnetic waves constantly withdraws energy from the oscillating circuit of the transmitter and therefore damps it, it is necessary constantly to supply fresh energy to it. This is done by means of a triode or a transistor in a feedback circuit (Figs. 1, 2 and 6). As a result of induction, a grid voltage is produced which supplies a correctly synchronised high-frequency alternating current to the oscillating circuit in the anode circuit or collector circuit. By means of a suitably connected microphone the transmitted wave is amplitude-modulated or frequency-modulated. In radio engineering "modulation" means that the pattern of the current variations or impulses are superimposed upon the carrier wave. With amplitude modulation it is the amplitude of the high-frequency carrier wave that is modified; on the other hand, with frequency modulation, the frequency of the carrier wave is varied in accordance with the frequency pattern of the sound waves (speech or music) to be transmitted. In the receiver the oscillating circuit is tuned to the frequency of the incoming wave, whose energy produces a synchronous high-frequency alternating current in the oscillating circuit. The alternating voltage associated with this current is used to control the anode current or collector current through the agency of the grid (Figs. 3 and 4) or the emitter (Fig. 5). The high-frequency alternating current component that is superimposed on the anode direct current is fed to a loudspeaker through a resistance or a transformer (Figs. 3, 4 and 5). The cases considered here relate to high-frequency amplification. If rectification is applied before amplification, the process is called low-frequency amplification. Cf. radio receiver (page 100), thermionic tube (page 90), and semiconductors (page 98).

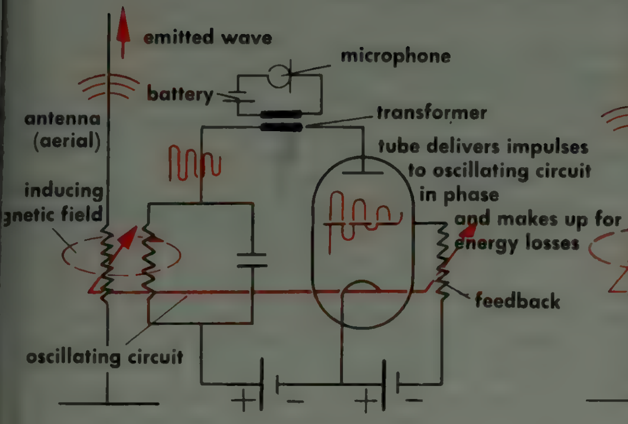


Fig. 1 TUBE TRANSMITTER
(*amplitude modulation*)

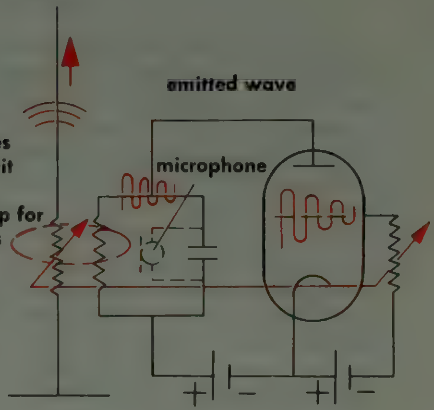


Fig. 2 TUBE TRANSMITTER
(*frequency modulation*)

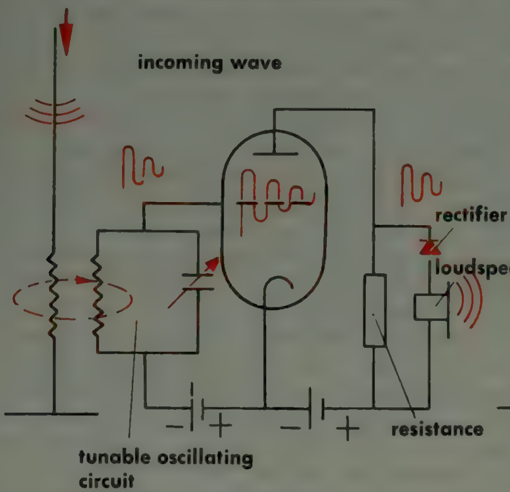


Fig. 3 TUBE RECEIVER
(*amplitude modulation*)

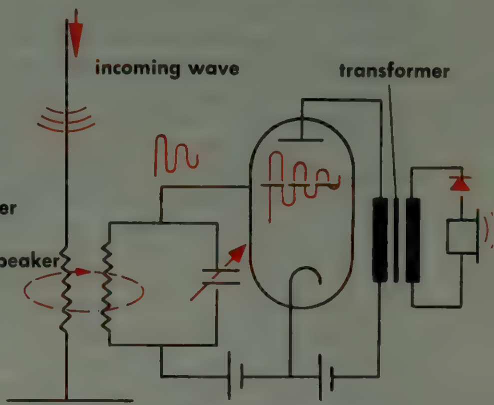


Fig. 4 TUBE RECEIVER WITH
TRANSFORMER

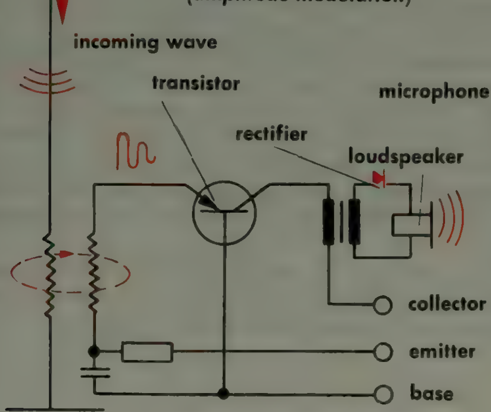


Fig. 5 TRANSISTOR RECEIVER

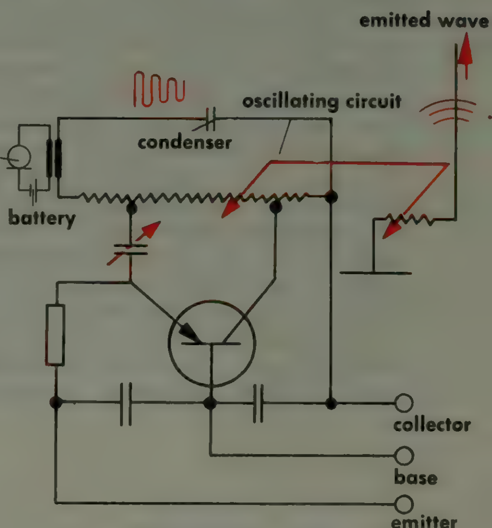


Fig. 6 TRANSISTOR TRANSMITTER

The "free" electrons which are present in the crystal lattice of a metal can be liberated by the addition of energy. The simplest way to do this is by heating the metal to incandescence. The emission of the electrons can in that case suitably be conceived as a kind of evaporation process (Fig. 1). As a result of collisions between the atoms of the metal—which are in a state of increased thermal agitation (heat motion) because of the heat applied—high-energy electrons are released, which can break out from the surface of the metal and escape despite the restraining effect of the attractive force exerted by the positive ions remaining behind in the metal. This emission of electrons from an incandescent metal (usually an electrically-heated filament) can most suitably be made to take place in a vacuum. This prevents oxidation of the very hot surface of the metal and allows the electrons to emerge unobstructed, i.e., without colliding with, or being neutralised by, gas molecules and ions of the air. A device which fulfils this requirement is the thermionic tube (or thermionic valve). It must contain at least two electrodes; also, it may have one or more additional electrodes for controlling the stream of free electrons flowing inside it. The simplest form, with only two electrodes, is the diode valve (or merely "diode") (Fig. 2). The hot cathode, connected to the negative pole of a battery, faces the anode (positive electrode) which draws electrons towards it from the cathode. On their way from the cathode to the anode the electrons form a negative space charge, which can be influenced by the electric fields of other electrodes (and also by magnetic fields; cf. magnetron, page 92). The simplest thermionic tube of this kind is the triode, which has a third electrode, the control grid. The latter allows the electrons to pass through it and controls them by appropriately modifying the space charge. The circuit connections for a triode of this kind are shown in Fig. 3. The effect of the grid upon the space charge is clarified in Fig. 4. Variations in the grid voltage strengthen or weaken the space charge and thus vary the density of the stream of electrons flowing to the anode. The cathode of the triode in Fig. 4 shows a particular feature in that the electrons are not emitted by the hot glowing filament itself, but by a coating of barium oxide which is (indirectly) heated by the filament. The oxide cathode offers two advantages: for one thing, it emits electrons already at a relatively low temperature (feeble red heat), so that the heat losses are lower; secondly, it forms an equipotential surface, so that the conditions of emission are the same for all the electrons. This latter condition is not satisfied if the electrons are emitted by the filament itself because a voltage drop occurs between its ends. The control performance of a triode can be judged from so-called characteristic curves. Each of these curves is obtained by measuring the current I_a flowing from the anode and also measuring the grid voltage U_g (this is the voltage which is applied to the grid at $\pm$ in Fig. 3) while the anode voltage U_a is kept constant at one particular value. One such curve can be plotted for each of a number of values of the anode voltage (U_{a1}, U_{a2}, U_{a3}). From the characteristic curves thus obtained (Fig. 5) it is possible to deduce certain characteristic quantities (e.g., the slope S) which can be used for describing the behaviour of the thermionic tube. By applying negative grid voltages, the flow of electrons can be completely inhibited; on the other hand, for high positive values of the grid voltage the electron flow must attain a maximum (saturation) value when the space charge has become exhausted and all the electrons emitted by the cathode reach the anode. The increase in the anode current from zero up to its saturation value is approximately linear.

The thermionic tube has the property of unipolar conductivity, i.e., the electrons can flow in one direction only, namely, from the (hot) cathode to the (cold) anode. For this reason, thermionic tubes can also serve as rectifiers (see page 96).

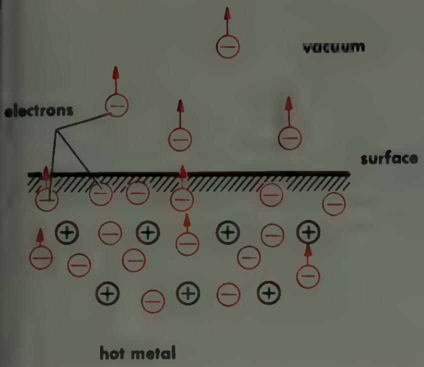


Fig. 1 EMISSION OF ELECTRONS FROM A HOT METAL

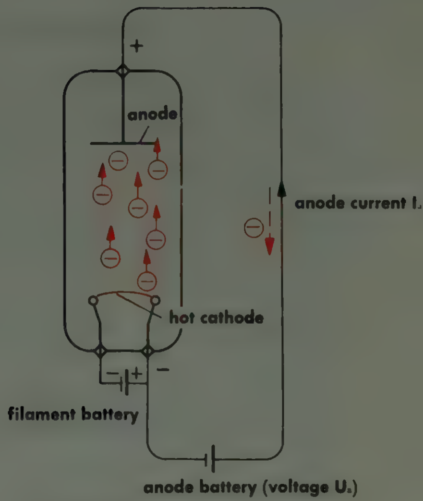


Fig. 2 DIODE

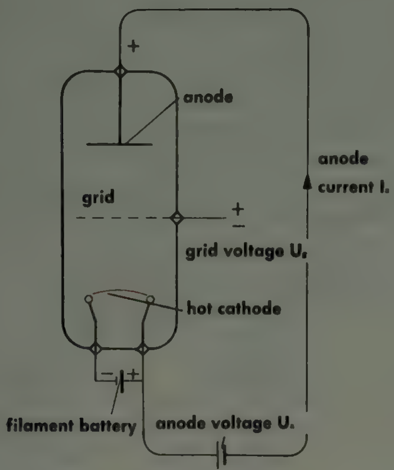


Fig. 3 TRIODE

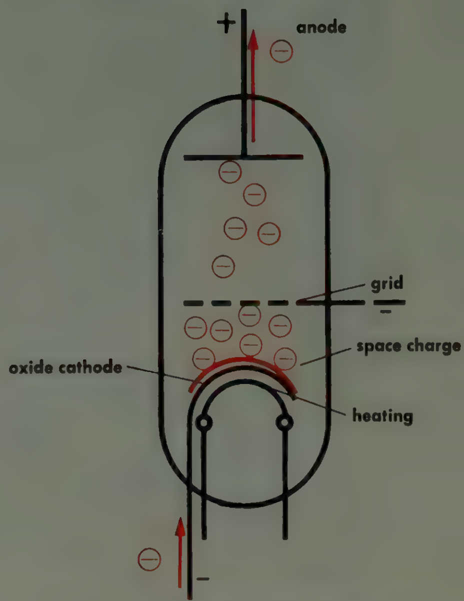


Fig. 4 EFFECT OF THE GRID ON THE SPACE CHARGE

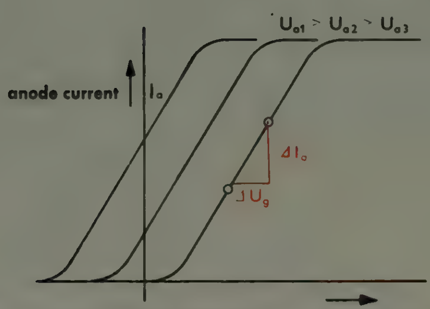
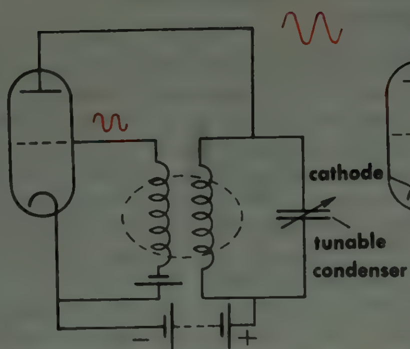


Fig. 5 TRIODE CHARACTERISTIC
STEEPNESS $S = \Delta I_a / \Delta U_g$

If variations occur in the electron density or the electron velocity of a stream of electrons, portions of the electromagnetic field which always surrounds the electron stream will become detached at the same rhythm, or frequency, as the variations and will—more particularly if the variations occur periodically—cause emission of electromagnetic vibrations. Such vibrations are thus invariably associated with the existence of a high-frequency alternating current. Conditions for the detachment of the electromagnetic field are more favourable according as the frequency is higher, because the individual phases will, on reversal, encounter an already changed situation (cf. page 82). The frequency of the alternating current which supplies the energy will depend on the inductance and capacitance (or capacity) of the circuit which generates and gives forth the electromagnetic energy. High frequencies, i.e., short wavelengths, will therefore be obtained more particularly by reducing the self-inductance and capacitance of the oscillating circuit and by using a set of apparatus that produces periodic variations of electron density and electron velocity (Figs. 1 and 2). Fig. 2 shows the so-called three-point circuit (1-2-3), in which the inductance of the terminal wires (L_A and L_G) provides the self-inductance and a very small condenser (C) merely serves to separate the anode and cathode. The self-inductances of the tubes serve as the capacitances of the oscillating circuits. If the inductance and capacitance are still further reduced (shorter terminal wires, smaller condenser), the finite transit time of the electrons will disturb the onset of the vibrations. If the transit time exceeds the oscillation period (which is the reciprocal value of the frequency) of the oscillating circuit, phase displacements will occur which may be likened to the behaviour of a swing which is being pushed at intervals of time which are shorter than its swinging period. To make a virtue of necessity, investigators either chose the transit time itself as the oscillation period (Barkhausen–Kurz oscillations; Fig. 3) or shortened the distances between the grid and the cathode, the electrodes being constructed as discs for the sake of better mechanical stability (disc-seal triodes; Fig. 4). In the first-mentioned case the grid must have a positive, but the anode a negative voltage, in order to make the electrons oscillate to and fro between the anode and cathode (and pass through the grid while doing so). The highest frequencies attainable with a triode are in the region of about 10^4 megahertz. The grid-to-cathode spacing of 15 microns (0.015 mm = 0.0006 inch) presents a technical limit to a further increase in frequency.

The next step in the process of development was to produce the periodic vibrations of the electron stream without the help of a material grid. A stream of electrons can be forced to perform a periodic motion other than oscillation, namely, movement along a circular or spiral path. This led to the construction of the *magnetron*. The permanent magnetic field with constant field strength functions as a virtual grid. The rotational frequency of the electrons on circular paths between anode and cathode determines the frequency of the electromagnetic vibrations (Figs. 5a–5c).

(Continued)



ig. 1 INDUCTIVE FEEDBACK CIRCUIT
(*Meissner circuit*)

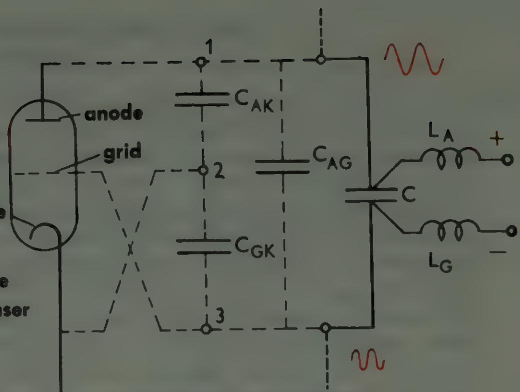
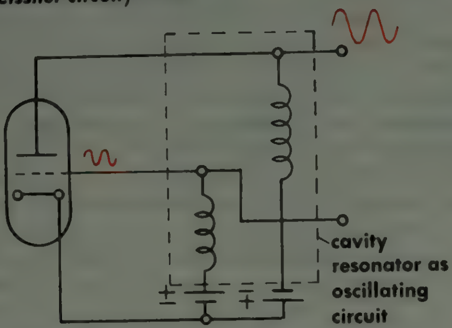


Fig. 2 THREE-POINT CIRCUIT FOR VERY HIGH FREQUENCIES



ig. 3 BARKHAUSEN-KURZ CIRCUIT WITH DISC-SEAL TRIODE FOR VERY HIGH FREQUENCIES

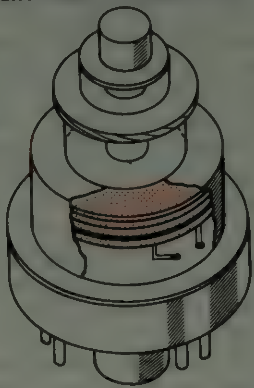


Fig. 4 DISC-SEAL TRIODE FOR OSCILLATORS

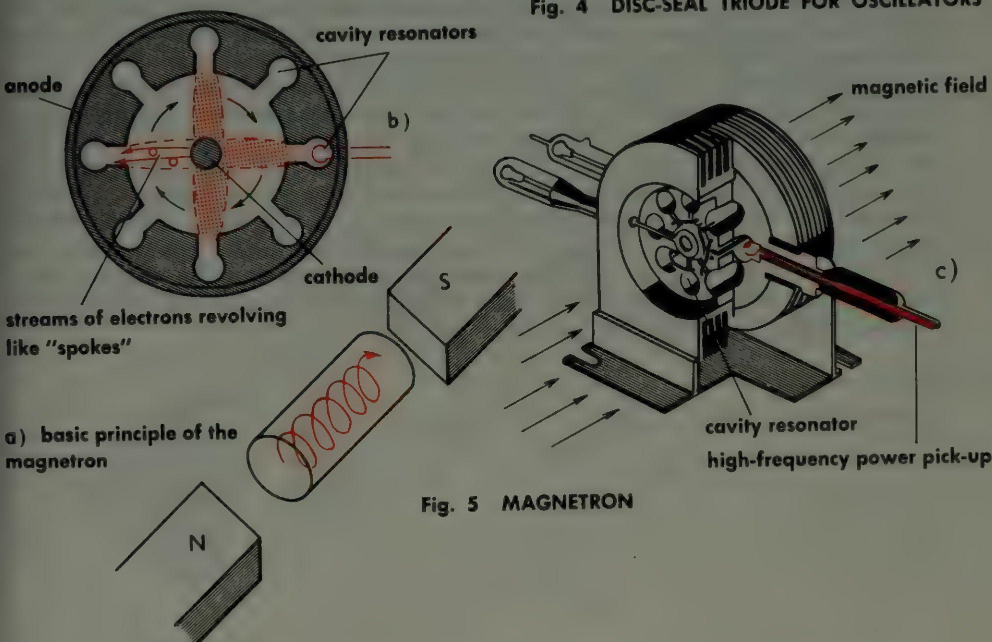


Fig. 5 MAGNETRON

Further increases in the frequencies attainable with the magnetron (see page 92) can be achieved by reducing the size of the operative components. There is obviously a technical limit to what can be done in this direction. In the arrangement shown in Fig. 1 an electron beam of constant velocity is directed into a cavity resonator (cf. page 86) which modulates the velocity of the electrons in the rhythm of the voltage impressed upon it. On the way from the cathode to the anode variations of density occur on account of the different electron transit times as a result of their different velocities, the electrons are periodically compressed to "electron packets". The intensified alternating current energy is abstracted from the beam in a second cavity resonator and fed to an aerial (antenna). Control is effected by feedback by means of a hollow waveguide to the input cavity resonator. A velocity-modulated tube of this kind is called a *klystron.* The *travelling-wave tube* (Fig. 2) utilises the fact that an electric field prefers a metallic path to a vacuum. If the metallic spiral which guides the field has a pitch of $\frac{1}{13}$, the field will have a velocity component whose magnitude in the flow direction of the electrons is equal to $\frac{1}{13}$ of the velocity of light and thus has about the same value as the electrons accelerated by about 1500 volts, which interact with the field. With electron tubes of this kind it is possible to attain ultra-high frequencies corresponding to wavelengths in the millimetre range. This already brings us into the range of long-wave heat radiation, i.e., the range of atomic phenomena. The next step, therefore, was to enable atomic electromagnetic radiation to be controlled at will. This can be achieved by means of the molecular oscillators (known as *masers* and *lasers* respectively.*) Masers differ from lasers in the wavelength range of the radiation emitted: it consists of microwaves in the case of the maser, and of light waves in the case of the laser. The principle is indicated in the atomic diagram in Fig. 3. Three possible different energy states I, II, III (stable electron orbits) occur around a positive atomic nucleus. The pump radiation (I) raises electrons from state I to state III, with the result that more electrons than the number corresponding to temperature equilibrium will be in state III. There they form a supply of electrons in readiness to move into state II, from where the electrons—under the action of an excitation radiation (2)—pass back to state I, this being accompanied by the emission of radiation (3).

The entire process (pumping, excitation, emission) takes place so rapidly that the equalisation due to thermal energy has not yet been accomplished. For this reason amplifiers and transmitters which operate on the maser or laser principle are practically noiseless, i.e., they do not exhibit the disturbances caused by the thermal oscillatory motion of the electrons (thermal noise; cf. page 140), which can otherwise be eliminated only by cooling to a temperature close to absolute zero (–273° C). With these devices it is possible to detect power outputs of as low as 10^{-28} Watts. Because of its atomic character, the maser or laser radiation has an extremely constant frequency and sharp concentration of the radiation into a beam. Fig. 4 shows the technical form of construction of a ruby maser.

* These names are abbreviations of "microwave amplification by stimulated emission of radiation" and "light amplification by stimulated emission of radiation".

Fig. 1 KLYSTRON

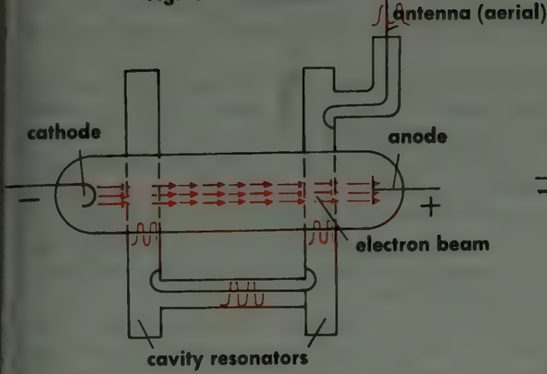
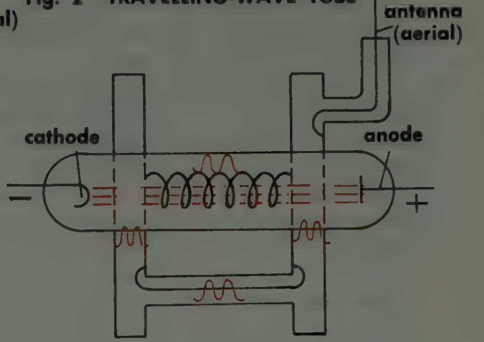


Fig. 2 TRAVELLING-WAVE TUBE



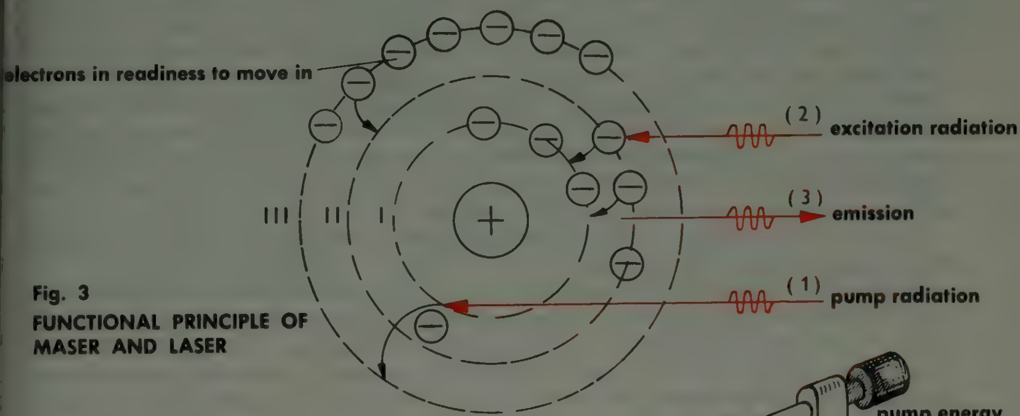


Fig. 3
FUNCTIONAL PRINCIPLE OF
MASER AND LASER

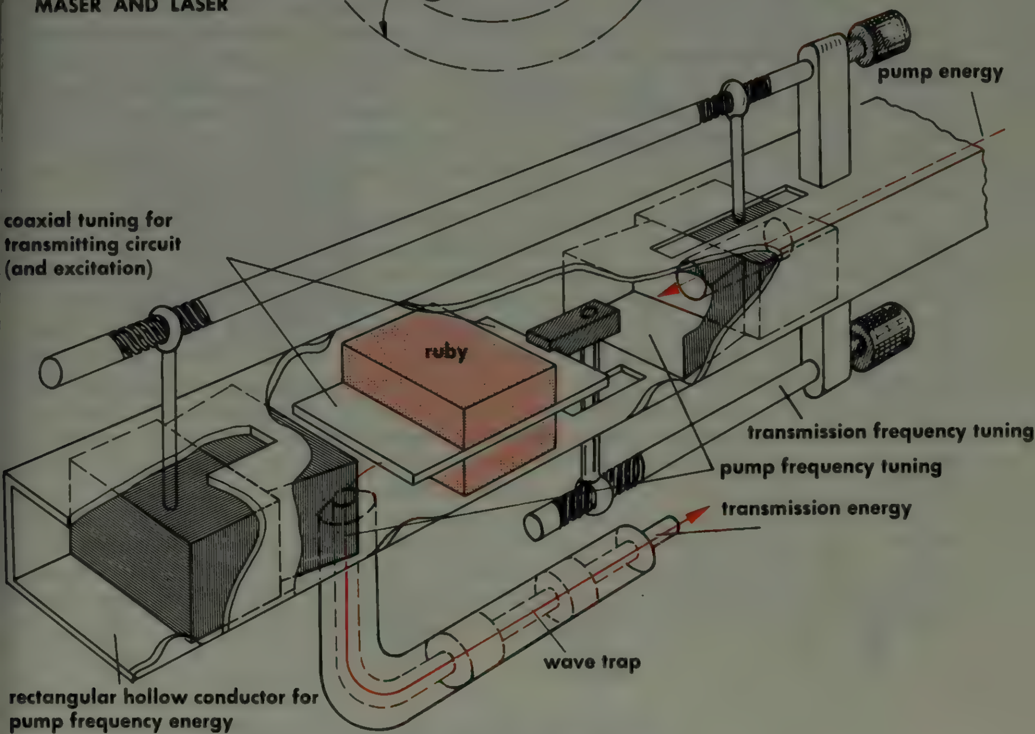


Fig. 4 RUBY MASER

A rectifier is a device which allows an electric current to pass in one direction but not in the opposite direction. A mechanical analogy is, for example, a venetian blind (Fig. 1).

A vacuum tube rectifier always has a cathode (K), which is almost invariably a filament heated by current, and an anode (A). The cathode emits electrons which either rush at high velocity straight to the anode and thus establish the conductive connection directly to the anode (in the case of high-vacuum tubes), or, alternatively, ionise the gas that is present in the tube (in gas-filled tubes). High-vacuum rectifier tubes can be used only for relatively small currents. Fig. 2 shows the principle of a rectifier in its simplest form (one-way rectifier): as the rectifier tube allows current to pass in one direction only (Fig. 3), a pulsating direct current is produced in the circuit; this current is "smoothed" by the condenser.

Gas-filled rectifier tubes (the gas filling usually consists of mercury vapour or an inert gas, e.g., argon or helium) differ in their functional behaviour from high-vacuum tubes. When a low anode voltage is applied, at first only a weak current will flow, as the gas retards rather than assists the movements of the emitted electrons. On increasing the voltage, however, so high a velocity is imparted to the electrons that they ionise the gas atoms as a result of colliding with them: a powerful current suddenly flows: the tube "ignites". The internal resistance of a conductive path through an ionised gas is substantially lower than that of an electron stream, and the power output is therefore higher. Fig. 4 shows the technical form of construction of a large rectifier. It is intended more particularly for three-phase alternating current and is accordingly provided with three anodes.

Besides the thermionic rectifiers described above, semiconductors (see page 98) are also used as rectifier elements.

A characteristic feature of semiconductors is that their low electric conductivity can be substantially modified by the addition of minute quantities of impurities. If the crystals are contaminated by the intentional addition of impurities, the conductivity is very greatly increased. For instance, one part in ten thousand of boron can increase the conductivity of silicon one million times. Depending on the kind of impurity, a positively conducting (*P*-type) or a negatively conducting (*N*-type) semiconductor can be obtained. The direction of flow of the electrons can be controlled in this way. If a *P*-type and an *N*-type semiconductor layer are in contact with each other in, for example, a germanium crystal, then such a crystal will have rectifying properties (Fig. 6; semiconductor diode).

If an external voltage is so applied to a semiconductor diode that the negative pole is in contact with the *N*-type layer and the positive pole is in contact with the *P*-type layer (Fig. 6a), then the negative charge carriers of the *N*-type layer and the positive charge carriers of the *P*-type layer are forced to the middle of the crystal—where the *P–N* boundary is located—because charges of similar sign repel each other. At the boundary an intensive interchange of electrons and "holes" (equivalent to positive electronic charges) occurs, so that a strong current flows. If the polarity of the externally applied voltage is reversed (Fig. 6b), the negative and positive charge carriers are drawn away from the boundary zone because charges of dissimilar sign attract each other. At the boundary—the so-called blocking layer—there are now no charge carriers available, and no current can flow.

Fig. 5 shows the characteristic of a semiconductor diode. The voltage and the current have been plotted in a co-ordinate system. It appears that the current increases when the voltage is raised towards positive values, but that the current remains zero when negative voltages are applied (though some current may flow when the negative voltage is further increased to a high value).

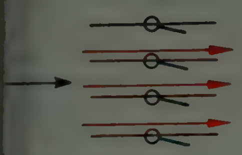


Fig. 1 VENETIAN BLIND AS "RECTIFIER"

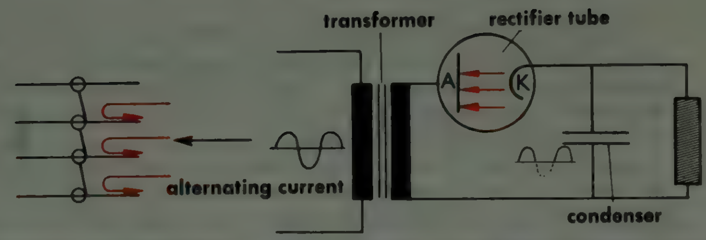


Fig. 2 SIMPLE ONE-WAY RECTIFIER TUBE

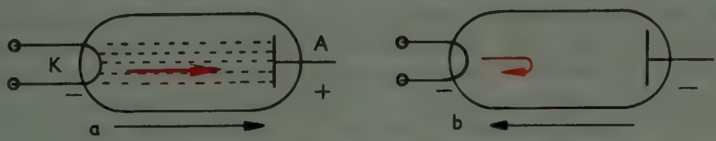


Fig. 3 DIAGRAM OF RECTIFIER TUBE: *current can flow in direction (a) but not in direction (b)*

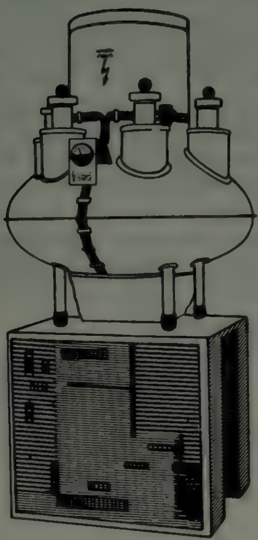


Fig. 4 A LARGE RECTIFIER

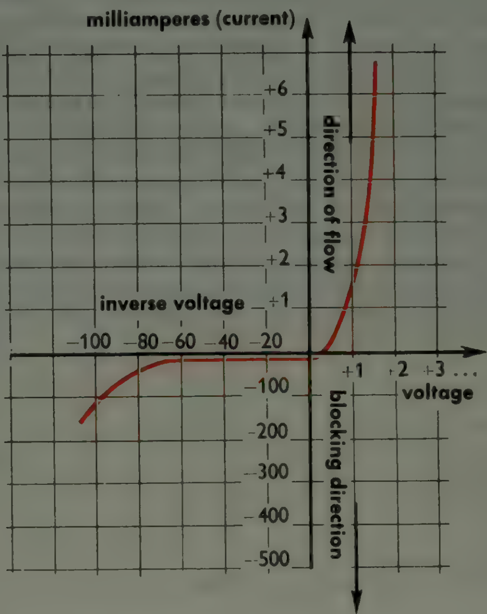


Fig. 5 CHARACTERISTIC OF A SEMI-CONDUCTOR DIODE

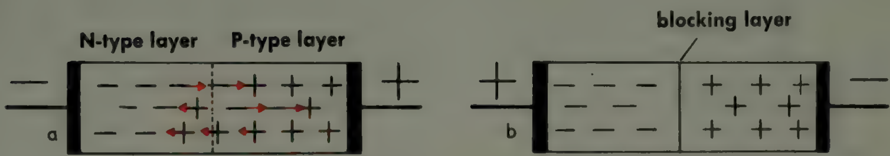


Fig. 6 RECTIFYING ACTION OF A SEMI-CONDUCTOR

In contrast with metals, semiconductors contain only a relatively small number of quasi-free electrons (see page 76) at room temperature. The conductivity of semiconductors is therefore of the order of 10^{-5} times as low as that of metals. It can be increased by breaking up the electronic double bonds which normally exist between the tetravelent atoms of semiconductors, as represented in the crystal lattice diagrams (Figs. 1 and 2). Commonly used semiconductors are germanium (Ge), silicon (Si) and various others. By the incorporation of impurities, i.e., foreign atoms of higher or lower valence, which function as points of disturbance in the crystal lattice (Fig. 4), the conductivity properties of semiconductors can be varied within very wide limits. A donor impurity makes electronic conduction possible by providing free electrons; an acceptor impurity captures electrons. Positive conduction occurs as the result of a deficiency of electrons: there is a flow of positive charges ("holes"); negative conduction occurs as the result of a surplus of electrons. Where positively and negatively conducting zones meet each other, the transition region constitutes a layer of high electric resistance (blocking layer) (Fig. 6). When a voltage is applied, the concentration of positive and negative charges in the transition region can be increased (flow direction of current) or decreased (blocking direction). Figs. 3 and 5 show the conditions within the so-called energy level diagram. The range ΔE is known as the "forbidden band". The ordinate is the electron energy E, the abscissa represents the longitudinal extension x. Semiconductor elements comprising a transition from positive to negative conduction are called *semiconductor diodes*, while those which comprise two such transitions are called *transistors*. The transistor (Fig. 7) is a semiconductor triode. Its three electrodes are respectively called the emitter, the base and the collector. The significant feature is that the base should be narrow (50 microns) so as to enable the charge carriers from the emitter to pass through the boundary layer 1 and traverse the base, so that they can thus affect the processes taking place at the boundary layer 2. Control of the passage of current between base and collector is thereby achieved. This enables the transistor to be employed for purposes of amplification and the generation of oscillations. It is progressively superseding the thermionic valves or tubes (see page 90). Besides taking up far less space, transistorised systems have the advantage of dispensing with the filament current. Cf. rectifiers, page 96.

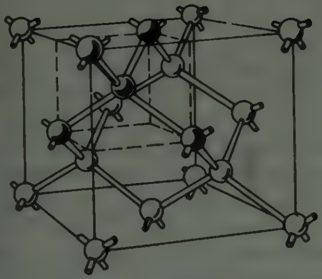


Fig. 1 CRYSTAL STRUCTURE OF A SEMI-CONDUCTOR

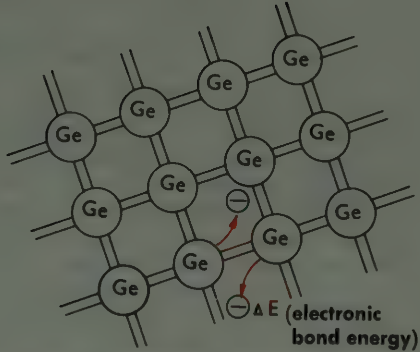


Fig. 2 CRYSTAL LATTICE DIAGRAM

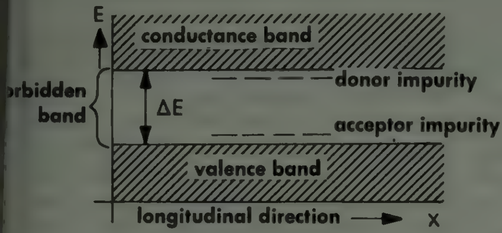


Fig. 3 ELECTRON BAND MODEL OF A SEMI-CONDUCTOR

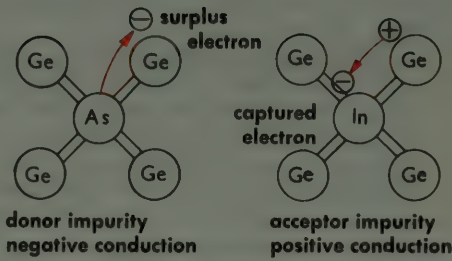


Fig. 4 DONOR AND ACCEPTOR IMPURITIES

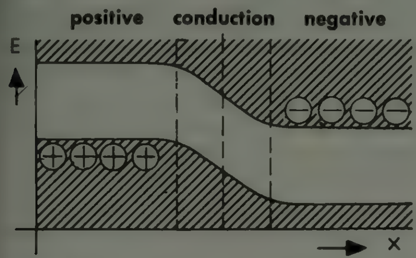


Fig. 5

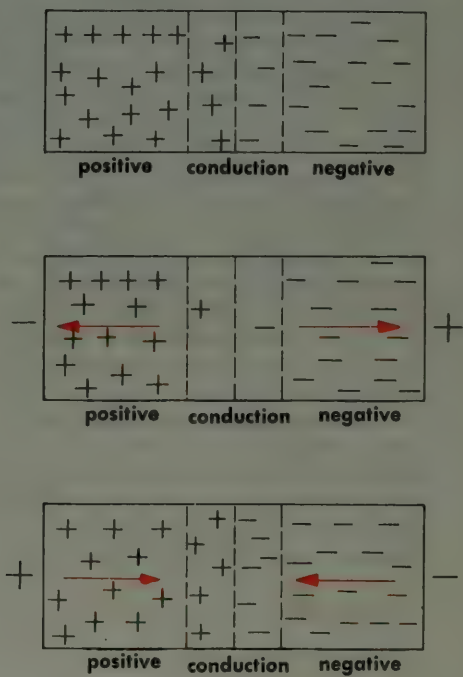


Fig. 6 FORMATION AND VARIATION OF TRANSITION REGION

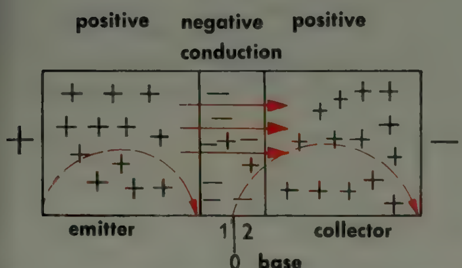


Fig. 7 HOW A TRANSISTOR WORKS: *current in collector circuit controlled by conduction in emitter circuit*

The radio receiver in its simplest form comprises an input circuit for tuning in to the frequencies of the various transmitters to be received, the demodulation circuit for separating the audio-frequency vibrations from the high-frequency carrier, a low-frequency amplifier stage, and the loudspeaker. The amplifier elements (high-vacuum tubes or transistors) are supplied with the necessary operating voltages by a suitable device. Corresponding to the frequency bands on which the various transmitters operate, receivers are equipped to receive long waves (150–285 kilohertz), medium waves (up to 1605 kilohertz), short waves (6 to 21.4 megahertz), and ultra-short waves (up to 100 megahertz). Long, medium and short wave reception function with a channel spacing of 9 kilohertz and with amplitude modulation (see page 88). The channel spacing in the ultra-short wave range is 300 kilohertz, and in this range frequency modulation is employed (see page 88). Reception here is usually better than in the other ranges, because high audio frequencies, which considerably affect the sound pattern, can also be transmitted. Besides, atmospheric disturbances, which have an amplitude-modulating effect, are of hardly any significance to this kind of reception.

The propagation conditions of the four radio waves ranges determine their possibilities of application and the purposes for which they are used. Long and medium wave transmitters send out a direct wave, which travels along the earth's surface, and an indirect wave. In general, only the direct wave is received (which has a range of up to some hundreds of miles); the indirect wave, which is reflected back from the Heaviside layer (an electrically conducting layer about 55 to 85 miles above the earth's surface), makes much larger ranges possible. In the short wave range, only indirect waves are used. Ultra-short waves are propagated in a quasi-optical manner (the waves travel in straight paths), so that they cannot travel beyond the optical horizon. To extend the range as much as possible, the transmitting antenna is installed at the top of a high mast or building.

The more tuning circuits that there are arranged in series in the radio receiver, the greater is the selectivity. In order to obviate manual tuning of the many circuits required, the superheterodyne receiver was developed (Fig. 1). A receiver of this kind comprises an oscillator which generates its own high-frequency oscillation, which is mixed with the oscillation received from the transmitter. The two circuits are so tuned by means of (rotary) variable condensers mounted on the same spindle[1] that the difference of the two frequencies is always the same. All the following circuits (intermediate frequency filters) have been pre-adjusted to this frequency difference by the manufacturer. In this way a high selectivity is obtained and yet only two circuits have to be tuned. Because of the considerably higher carrier frequency in the ultra-short wave range, receivers for this range require small tuning circuits which are usually installed as complete assemblies. In general, however, the variable condensers are mounted on the same spindle, like those used in medium wave receivers (cf. Fig. 2).

The low-frequency or audio-frequency part of the receiver is generally separated into a number of channels for high-pitched and low-pitched tones. Also, separate speakers for high and low notes are employed. In addition to the equipment described here, there is of course a wide range of special receiving equipment designed for the particular purpose for which it is intended. Cf. Alternating current, three-phase current. electromagnetic waves. page 82 *et seq.*

1. Shaft in U.S.A.

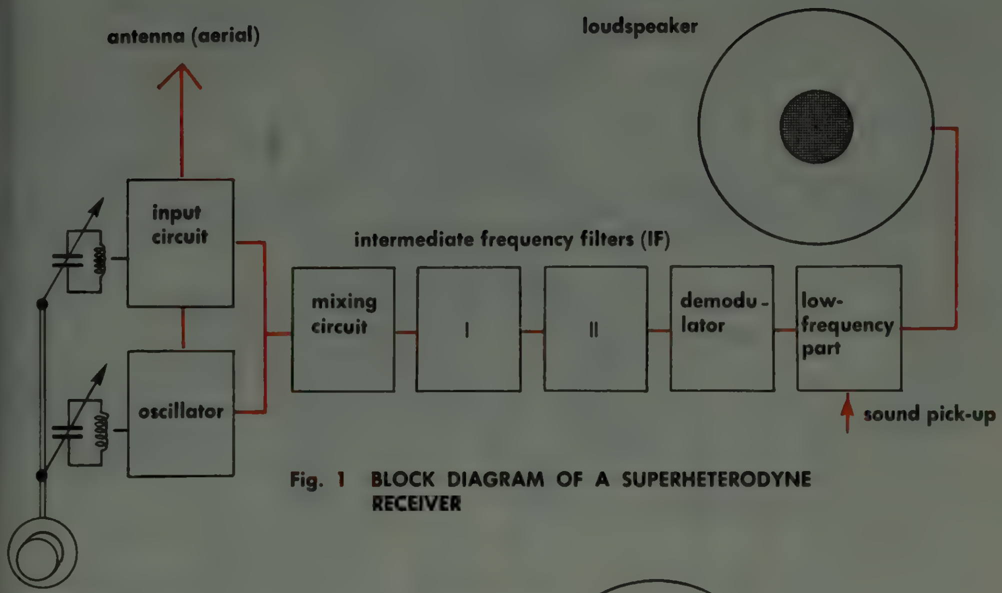


Fig. 1 BLOCK DIAGRAM OF A SUPERHETERODYNE RECEIVER

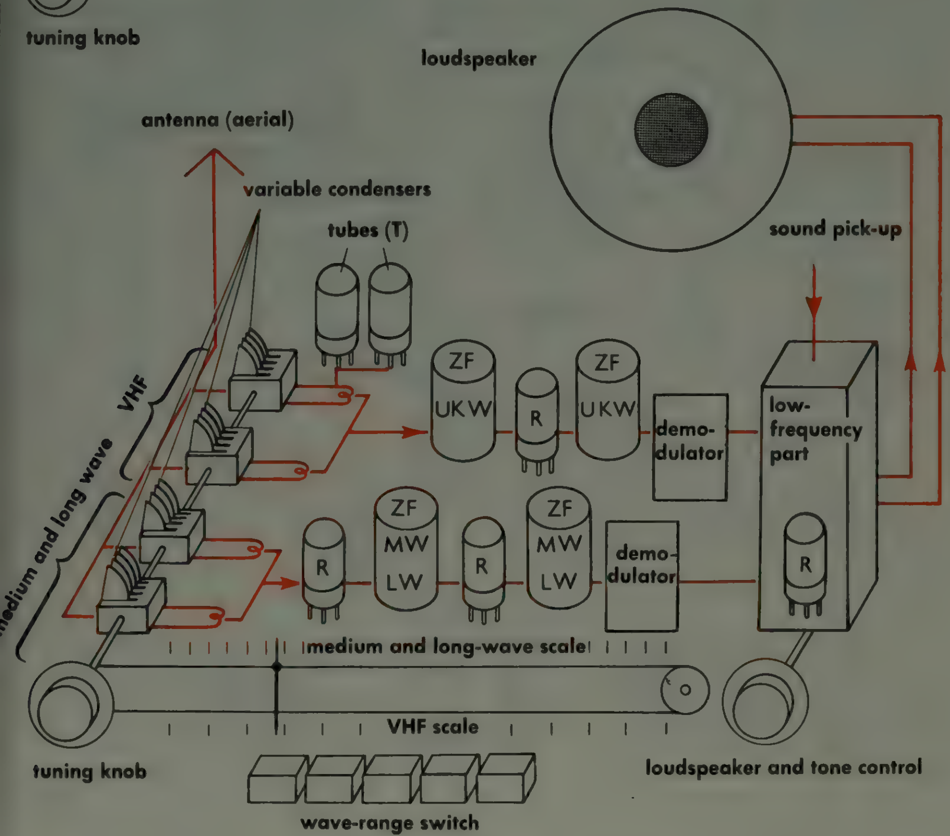
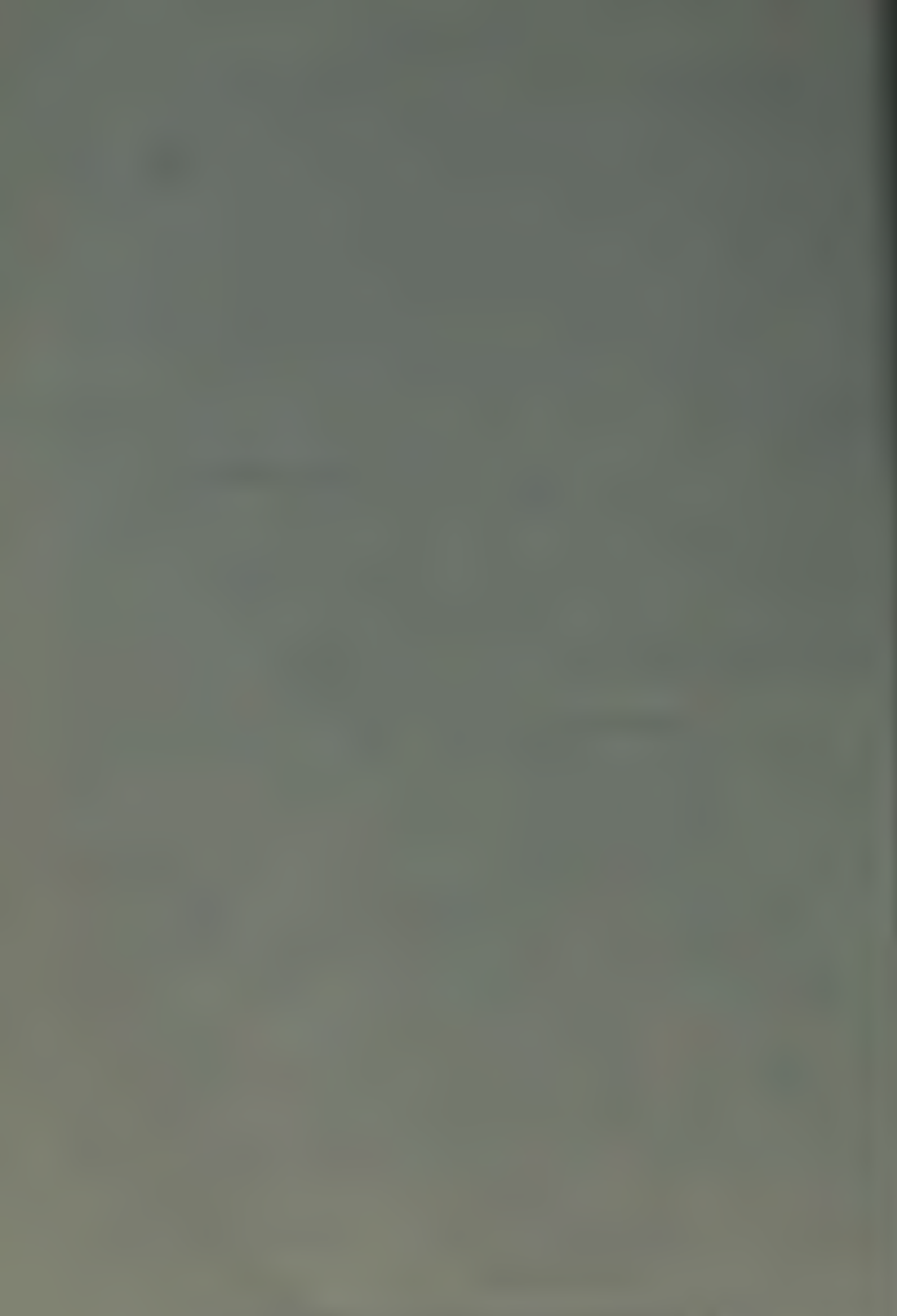


Fig. 2 DIAGRAM SHOWING MAIN COMPONENTS OF A RADIO RECEIVER



Receiver mounting showing input circuit, Albiswerk Zurich, Switzerland
Photo Roland Schneider, Len Sirman Press

A loudspeaker is a device for converting variations of electric energy into corresponding variations of acoustic energy, i.e., sound. Its task is therefore similar to that of a telephone receiver (cf. page 130), except that the sound produced is much louder. In fact, the early loudspeakers were designed like large telephone receivers (Fig. 1): mounted in front of the poles of a permanent magnet whose field is strengthened and weakened by the speaker current passing through coils is a metal diaphragm which vibrates to the rhythm of the field strength variations and transmits these vibrations to the air as sound waves. To improve the effect, a conical horn is fitted in front of the diaphragm. Because of the restraint of the diaphragm at its edges, where it is gripped in its mounting, the fidelity of the reproduction is adversely affected, however. The further development of the loudspeaker therefore had to aim at achieving, as far as possible, unrestrained vibration of the diaphragm. The first loudspeakers in which this principle was applied were constructed as shown in Fig. 2: the "diaphragm" is a resiliently mounted paper cone which is set in motion by the armature which is energised by the speaker current which here, too, can vibrate freely in the field of a permanent magnet.

A further advance is represented by the dynamic loudspeakers (also known as moving-coil loudspeakers). In such speakers the "armature" which vibrates in the magnetic field consists of a coil attached to the conical diaphragm. In the electrodynamic speaker (Fig. 3) the moving coil oscillates inside an electromagnet which is energised with direct current, while in the permanent-magnet moving-coil speaker (Fig. 4) the coil oscillates in an annular cavity of a specially-shaped permanent magnet.

All the loudspeakers described above use the electrodynamic principle for the conversion of electrical oscillations into mechanical vibrations which in turn produce sound waves in the air. Crystal loudspeakers (Fig. 5) and electrostatic loudspeakers (Fig. 6) are based on different principles. The crystal loudspeaker utilises the piezoelectric effect, i.e., the phenomenon that certain crystals (quartz, Seignette salt) develop an electric charge or potential difference when subjected to mechanical pressure and conversely undergo changes in thickness (and thus produce mechanical forces) when they are electrically charged by the application of a potential difference. Thus, when an alternating voltage is applied, the crystal undergoes periodic variations in thickness, i.e., thickness oscillations, which are transmitted to the loudspeaker diaphragm (Fig. 5). The electrostatic loudspeaker makes use of the electrostatic attractive and repulsive forces to which a diaphragm is subjected in the electric field of a condenser when the voltage applied to the latter is made to vary. The condenser plates are perforated, so that the sound waves can emerge through them. The two last-mentioned types of loudspeaker are more particularly suitable for the reproduction of high frequencies. In high-fidelity ("hi-fi") systems these speakers are used in combination with electrodynamic speakers to obtain sound-reproduction with a very high degree of accuracy (cf. vol. II, page 14 *et seq.*).

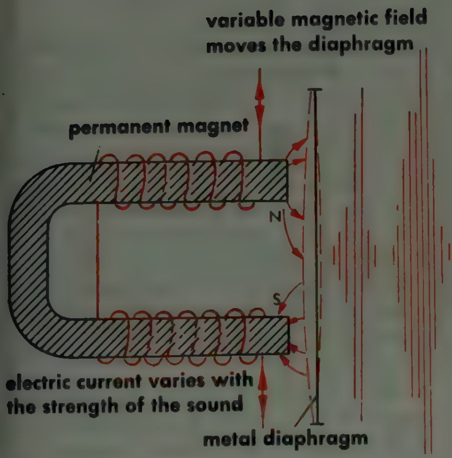


Fig. 1 PRINCIPLE OF LOUDSPEAKER

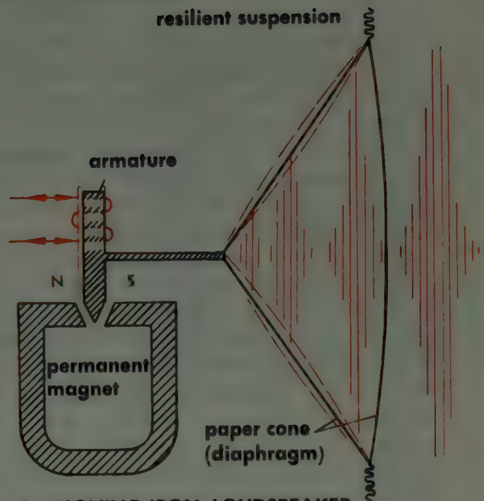


Fig. 2 MOVING-IRON LOUDSPEAKER

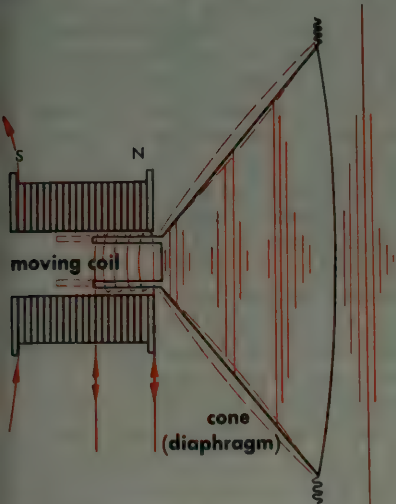


Fig. 3 DYNAMIC LOUDSPEAKER

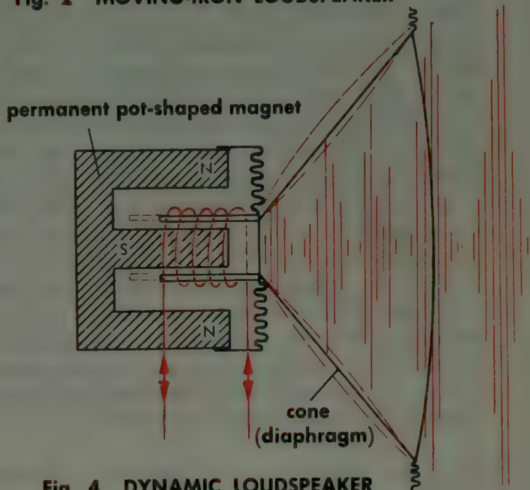


Fig. 4 DYNAMIC LOUDSPEAKER

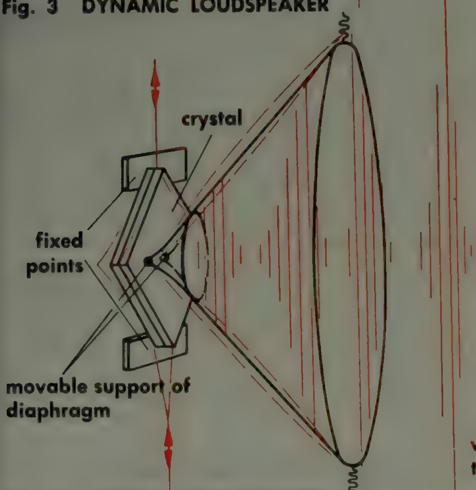


Fig. 5 CRYSTAL LOUDSPEAKER

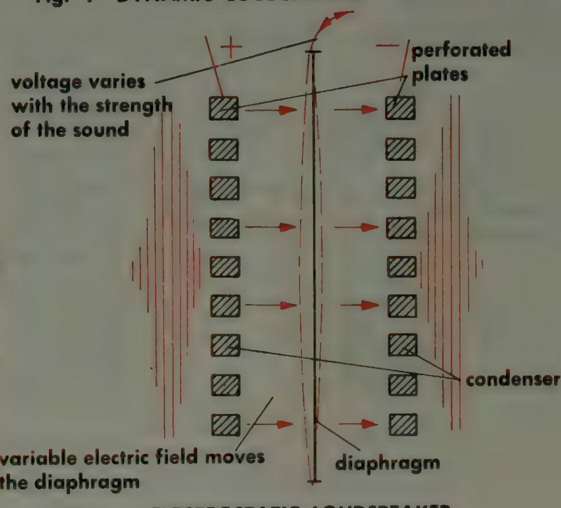


Fig. 6 ELECTROSTATIC LOUDSPEAKER

The electric power from the generating stations, which are situated in regions where coal or water power is readily available, is distributed by long-distance transmission lines. These are usually of the overhead type consisting of conductors ("wires") suspended from lattice steel towers. According to Ohm's law ($U = R.I$, where U is the voltage, R is the resistance, and I is the current strength) the voltage drop is greater in proportion as the resistance (i.e., the length) of the line is greater and the current strength (amperage) is higher. For this reason it is endeavoured to keep these two electrical quantities as small as possible. With regard to reducing the resistance, one is soon up against a limit, because for reasons of economy and weight it is of course not possible to increase the thickness of the conductors indefinitely. As for the current, however, this can be considerably reduced by using alternating current in conjunction with transformation (see page 110). The voltage drop along the transmission line is thereby also reduced.

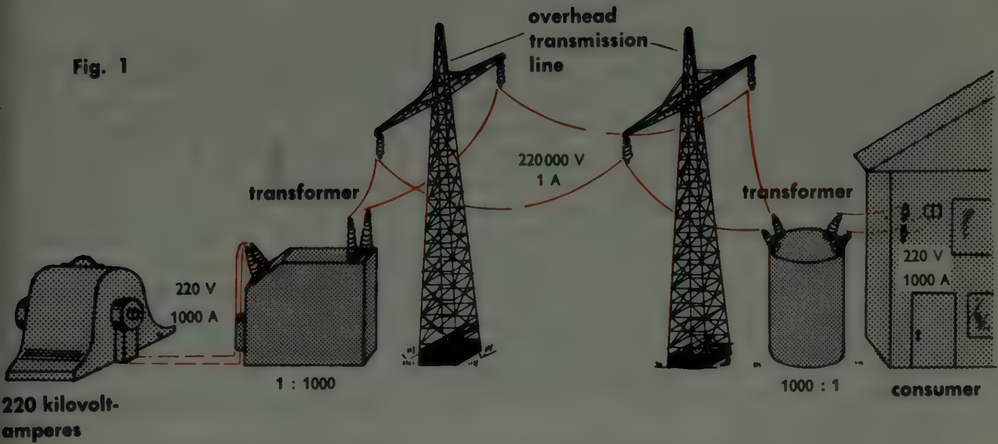
A simple example is illustrated in Fig. 1: a 220 kilowatt generator supplies a current of 1000 A at 220 V to the primary circuit of a high-tension transformer which increases the voltage a thousandfold and whose 220 kilowatt output is transmitted through a long-distance power line as a current of 1 A at 220,000 V. Before the electric power is supplied to the consumer, its voltage is transformed down in the ratio of 1000 to 1, so that a current of 1000 A at 220 V is again obtained. In actual practice the high tension of 220,000 V is usually first transformed down to 20,000 or 6000 V as an intermediate stage, which is used for local distribution lines, the final "step-down" of the voltage to 220 V[1] being performed in a second transformation stage at or near the consumer's premises. Since the resistance of a power line to alternating current is higher than its resistance to direct current, the use of alternating current involves additional losses. Such losses could be avoided by using high-tension direct current, but this calls for dependable high-duty rectifier equipment (see page 96). Long-distance power transmission with direct current is in fact being tried out experimentally in various countries.

The high-tension power transmission lines of the various generating stations are interconnected in a network known as the "grid". A system of this kind, of course, requires elaborate switchgear. High-tension switchgear and transformer equipment is very often installed (suitably sealed against moisture, etc.) in the open air.

Fig. 3 schematically illustrates the distribution of electricity by means of transmission lines and transformers from the power station to the individual consumers.

1. Consumer transmission lines carry 120 volts in U.S.A.

Fig. 1



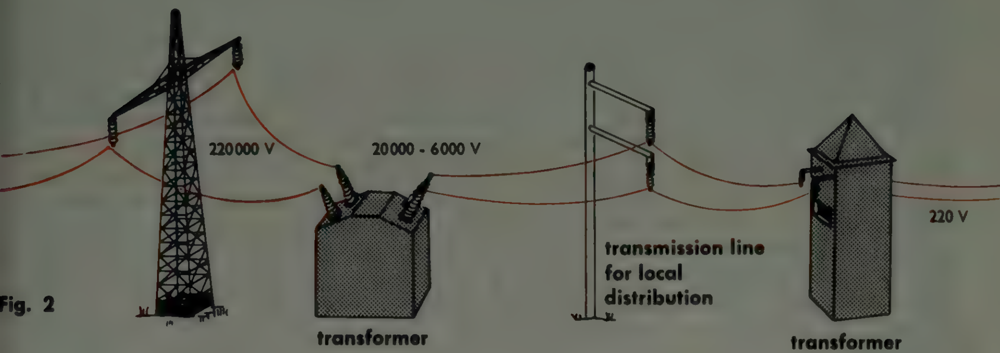


Fig. 2

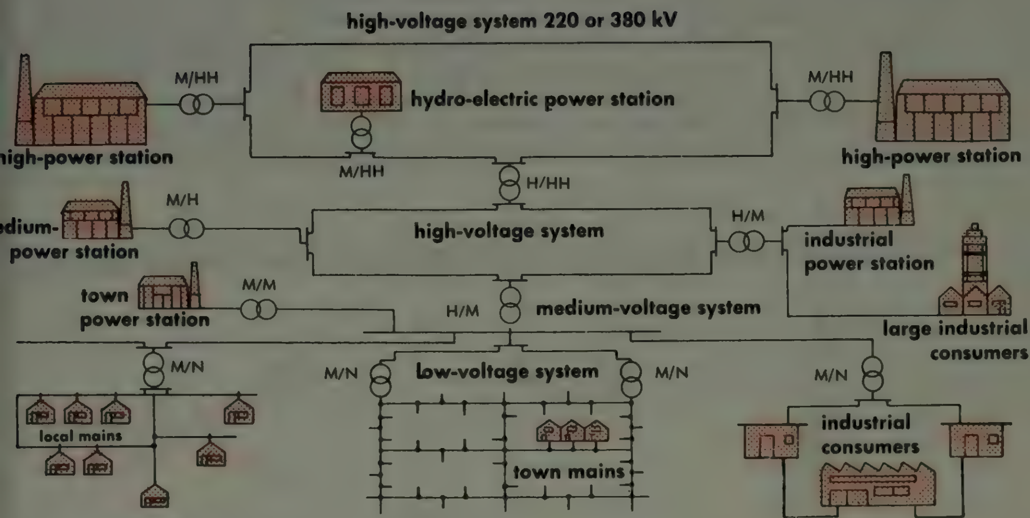


Fig. 3

HH = very high voltage   H = high voltage   M = medium voltage   N = low voltage

/ = is transformed to   ϴ = transformation

The substances of which the earth's crust consists mostly have a moderate degree of electric conductivity. Besides, the water which is often abundantly present in the soil contains salts in solution and thus forms an electrolyte which conducts electricity fairly well. This means that electric currents can pass through the soil if differences in potential (voltage) occur at various points of the earth's surface. Conversely, any potential difference will be immediately equalised, so that this surface in effect constitutes an equipotential surface (i.e., with a potential which is the same at all points). Since measurements of potential are merely relative, the earth's potential is adopted as a reference value and is, for the sake of convenience, taken as zero. Electrical equipment or conductors which are connected to the earth are said to be "earthed" or "grounded" (Figs. 1 and 2), which means that no difference of potential in relation to the earth can occur in them. Earthing thus provides a safeguard against electric shocks to anyone who happens to touch the metal parts concerned. A person who touches a "live" metal part which is not earthed is liable to receive a shock (possibly a fatal one).

Earthing can also be utilised for obtaining a "field-free" or "zero field" space, i.e., a space in which there is no electric field that might, for example, disturb delicate electrical measurements. For this purpose the walls of the laboratory are lined with wire netting or metal plates (and the doors and windows are also covered in this way) which are electrically interconnected and earthed. The room is then enclosed in equipotential surfaces which have the same potential as the earth, i.e., zero potential, so that no potential gradient (differences in potential from one point to another) and therefore no electric field can develop inside it. Faraday was the first to apply this principle, and for this reason a space screened against external fields is called a Faraday cage (in particular, this term is applied to an earthed wire cage surrounding apparatus to be protected from outside influence) (Fig. 3). The "pollution" of the atmosphere with electromagnetic fields from a multitude of radio and television transmitters has made Faraday's discovery a particularly important one for various present-day scientific and technical purposes. The all-metal body of a modern motor car also forms a kind of Faraday cage and provides excellent protection for the occupants against external electrostatic influences (lightning, in particular); besides, present-day motor tyres contain carbon black which makes the rubber conductive to electricity and thus ensures adequate earthing. The prerequisite for effective earthing is that the contact resistance between the earth wire and the soil should be very low, i.e., very good contact should be provided. Various methods of establishing a good earth connection are illustrated in Figs. 4a–4c: wide metal strips (4a) or metal rods (4c) well spread out and buried deep in the ground; connection to an underground metallic pipe system (e.g., water supply pipes) (4b).

Fig. 1 ELECTRIC APPLIANCE EARTHED BY MEANS OF A PROTECTIVE PLUG

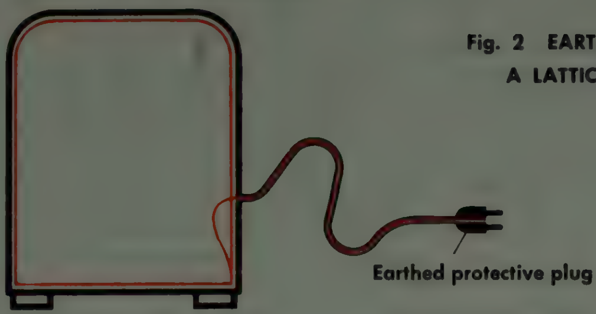


Fig. 2 EARTHING OF A LATTICE PYLON

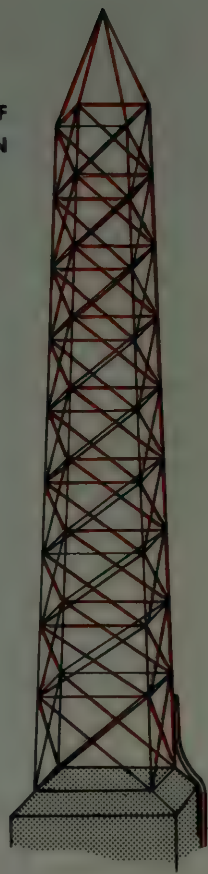


Fig. 3 FARADAY CAGE (*field-free space inside it*)

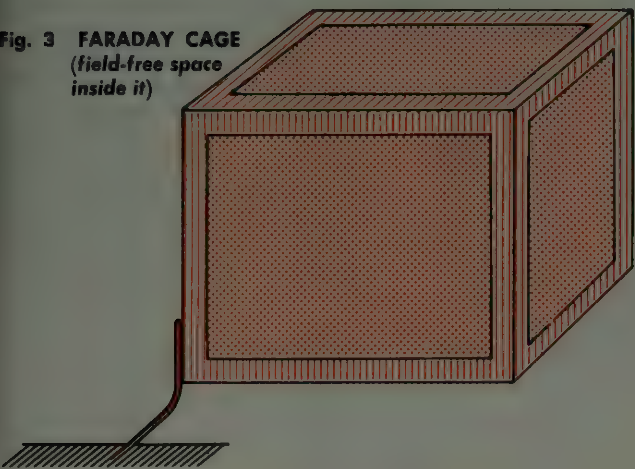
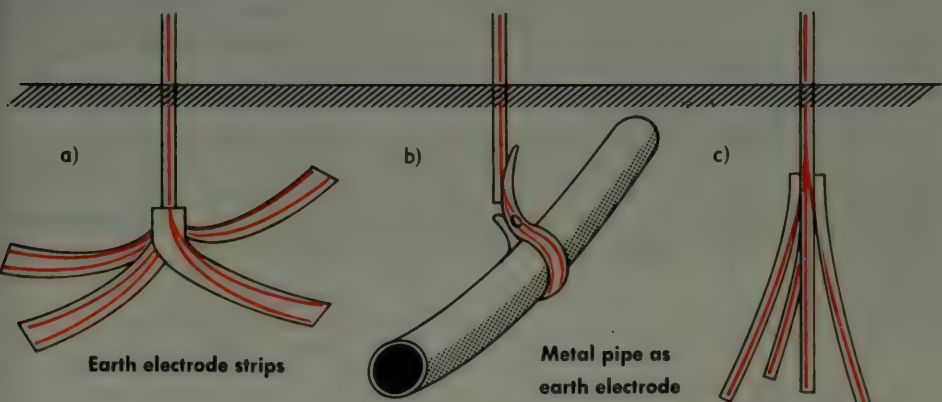


Fig. 4 EARTH CONNECTION



Transformers are used for "stepping up" alternating current (and three-phase current) to high voltages for long-distance power transmission—in order to minimise the relative voltage losses—and also for "stepping down" the voltage at the point of consumption (see page 106).

The principle underlying the operation of a transformer is closely associated with Faraday's law of induction, which states that when the magnetic flux enclosed within a circuit varies, an electric current which is proportional to the rate of variation is induced in the circuit. If the magnetic flux variation is produced by means of an electromagnet coil energised by alternating (primary winding of the transformer), an induced electric current can be obtained from a second coil (secondary winding of the transformer) through which the varying magnetic flux from the first coil is made to pass. The two coils, i.e., the primary and the secondary winding, are mounted on the same iron core, so as to obtain maximum concentration of the flux. Transformation is characterised by its very high efficiency (98–99%), practically the only energy losses associated with low-frequency alternating current being heat losses, namely, heat developed in the resistance of the copper windings (copper losses) and heat due to eddy currents set up in the iron core (iron losses). To minimise the eddy currents, the core is laminated, i.e., it is made up of a number of thin iron plates which are insulated from one another. The iron losses additionally comprise so-called hysteresis losses. Ideally: $U_1.I_1 = U_2.I_2$ (input voltage × input current = output voltage × output current). The ratio $U_1:U_2$ is determined solely by the ratio of the number of turns of the primary and secondary winding respectively; it is called the transformation ratio. Thus, if the secondary winding has twice as many turns as the primary winding, the output voltage will be twice as high as the input voltage, but the output current will be only half the input current (see Fig. 1).

To step up (increase) the voltage, the secondary winding of the transformer has a larger number of turns of wire than the primary winding. To step down (decrease) the voltage before the current is supplied to the consumer, a transformer is used whose primary winding contains a larger number of turns than the secondary winding. For local distribution in a town the high voltage of the power grid is transformed down to, say, 6000 volts and this in turn is reduced to 220[1] volts for distribution to individual consumers within the limited area served by the final step-down transformer. A low-voltage device such as an electric bell (see page 290) often obtains its electricity from the mains through a small transformer of its own (called a bell transformer) which gives current at 4–8 volts. Such a transformer is often of the kind known as an auto-transformer, which has only one winding, part of which forms the secondary, while the whole forms the primary, or vice versa. Its principle is illustrated in Figs. 3a and 3b. It is cheaper in construction than a transformer with two windings, and the copper losses are lower. The ignition coil of an internal combustion engine operates on the same principle, except that it is a step-up transformer which has to deliver a high voltage for the sparking plugs. In this case the low (primary) voltage is applied to part of the winding and the high (secondary) voltage is taken from the whole winding (see page 206, vol. II).

A commonly employed type of transformer is shown in Fig. 2. It is wound in concentric form, the primary winding being within the secondary. The iron core forms a double closed magnetic circuit (and is laminated, as already described). For practical purposes this is a more efficient design than the so-called toroidal transformer in Fig. 1, which merely serves to illustrate the principle.

1. 110/120 volts in U.S.A.

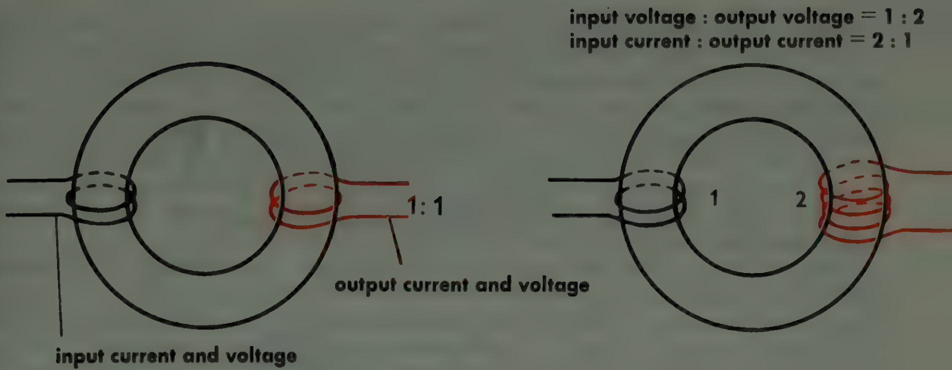


Fig. 1 TOROIDAL TRANSFORMER

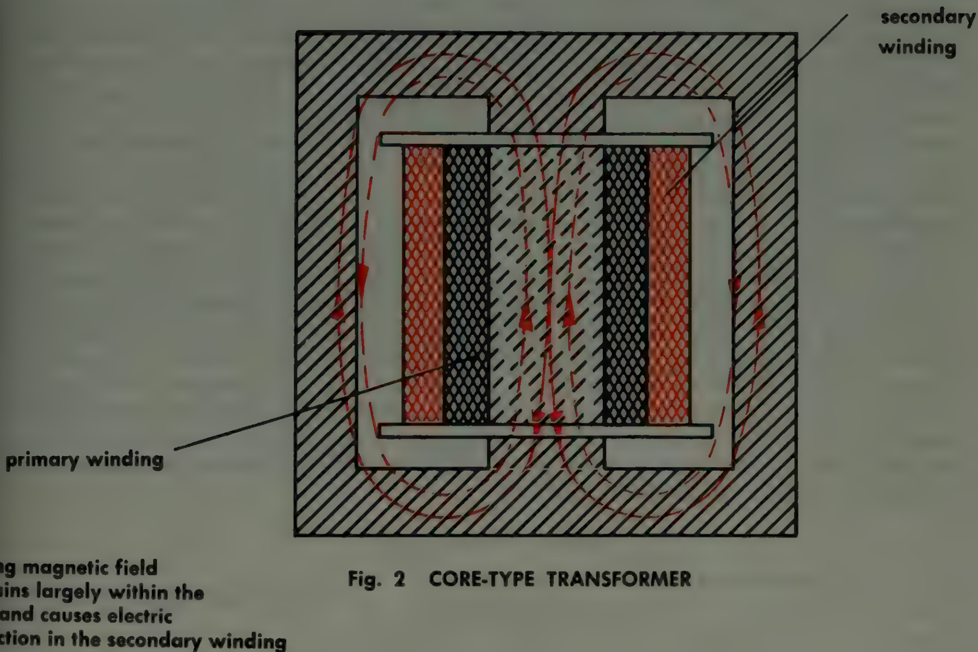


Fig. 2 CORE-TYPE TRANSFORMER

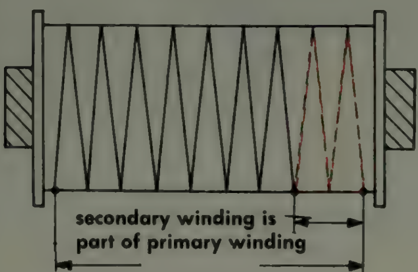


Fig. 3a AUTO-TRANSFORMER

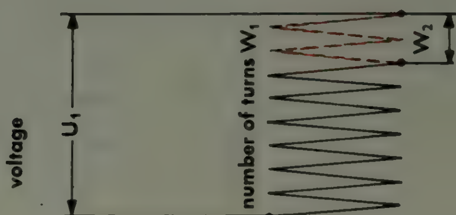


Fig. 3b DIAGRAM SHOWING PRINCIPLE OF AUTO-TRANSFORMER

A relay is an electric switching device comprising one or more contacts which open or close circuits. The switching device is mostly actuated by an electromagnet which closes or opens the contacts by means of a movable armature which it attracts or releases. However, there are also relays which are operated by other than electromagnetic forces, e.g., electrical attraction forces or mechanical forces such as the flexural force of a bimetallic strip in a thermo-relay (cf. thermostat, page 26).

In this article only the electromagnetic relay will be described. In a variety of forms it plays a very important part in many branches of electrical engineering (telecommunication, automatic control, computers, etc.).

Three electromagnetic relays differing in the design of the armature are illustrated in Figs. 1–3. Each relay has a coil of wire with an iron core and an iron yoke which carries the movable armature (or may be an integral feature of the latter). The yoke, which serves as an easy path for the magnetic flux, imparts the polarity of the rear end of the core to the armature, which is thus powerfully attracted by the opposite polarity of the front end of the core. To prevent the armature from remaining sticking to the core by the action of remanent magnetism (the residual magnetism which remains in the core even when no current is flowing in the coil), a small separator stud made of a non-magnetic material (brass) maintains an air gap between the armature and the electromagnet core. The current for energising the coil is supplied through the connections designated as "soldering lugs". In the relay in Fig. 1 the contacts are normally open: when the relay coil is energised, the core attracts the armature, which presses the bottom contact up and thus closes the contacts, so that current then flows through the working circuit by way of the connections 1 and 2. Several sets of contacts can be installed in a relay, as in Figs. 2 and 3; these are simultaneously actuated when the relay is energised. These include normally closed contacts, which open only when the relay is energised and then break the working circuit in which they are installed (connections 2 and 3). A type of relay which is used particularly in telegraphy is the polarised relay (Fig. 4). The armature, which carries the contacts at its front end, is suspended from a torsion wire and receives the polarity of a magnetic north pole from the suitably mounted permanent magnet. The rear end of the armature extends into a gap in an iron yoke with magnetic south polarity. Mounted on this yoke is the relay coil which produces the controlling magnetic flux. The superposition of the magnetic fluxes, and therefore of the forces exerted, is shown in Fig. 4. A relay of this kind responds differently to energising currents flowing in different directions.

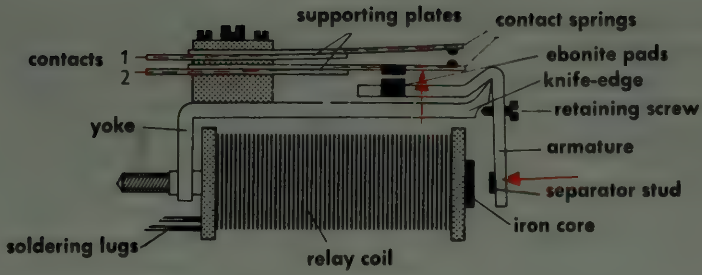


Fig. 1

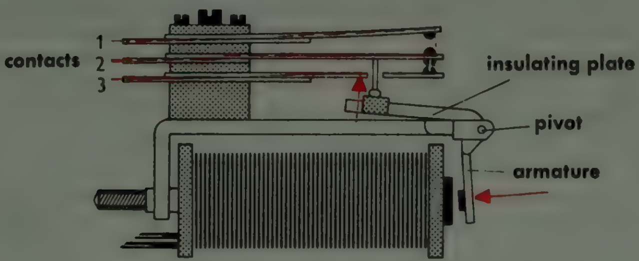


Fig. 2

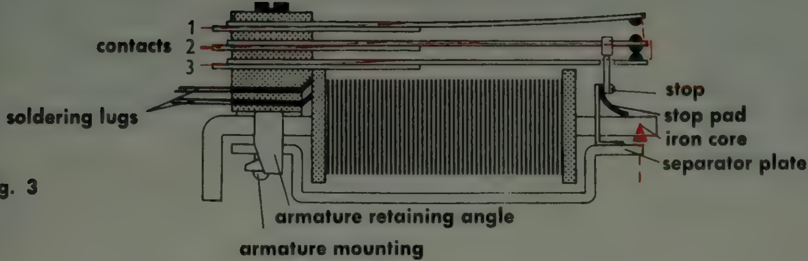


Fig. 3

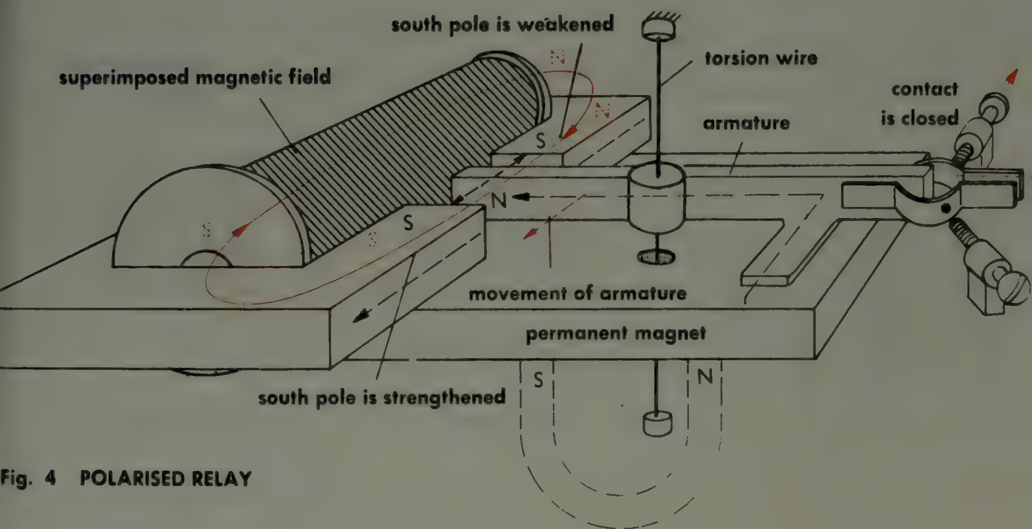


Fig. 4 POLARISED RELAY

A switch is a general name for a device used for effecting the completion and interruption of an electric circuit. Fig. 1 shows a very simple switch for low voltage current. This type is extensively used in telecommunication engineering and for many other purposes, e.g., in bell circuits. The switch lever is conductively connected to the contact *K* and can be connected to either of two alternative circuits by setting it in the position I or II. A switch of this kind is unsuitable for heavy currents because considerable sparking would occur, more particularly on interruption of the current, and this could give rise to dangerous arc formation. To prevent this, the interruption of the current must be accomplished as speedily as possible. For this reason knife switches are used. One such switch is shown in Fig. 2. The switch blade is connected to the lever by a spring. When the operating lever is pulled out, the spring is tensioned and then quickly pulls the blade out of the contact: the brief interruption spark then cannot develop into an arc. Another way to prevent arcing at the contacts is embodied in the mercury switch, which is filled with a protective gas (inert gas, nitrogen) (Fig. 3). The switches used for domestic lighting purposes are turn switches, tumbler switches or push-button switches (Figs. 4–6). In all three types the interruption of the current is effected suddenly, by a kind of spring operated trigger action which ensures rapid separation of the contacts, so that no harmful arcing occurs. Switches for very strong currents are usually of the electromagnetically operated type (Fig. 7). High tension switchgear is often of the oil break type, i.e., the contacts are separated under oil for the quick and effective extinction of the arc; the oil also serves as insulation (Fig. 8). Switches for very high voltages are often of the gas-blast type, in which a blast of high pressure hydrogen, air or other gas is directed on to the arc at the moment of separation of contacts to accelerate its extinction (Fig. 9).

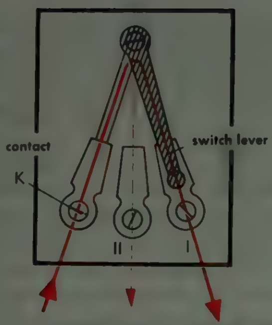


Fig. 1 SWITCH FOR LOW-VOLTAGE CURRENT

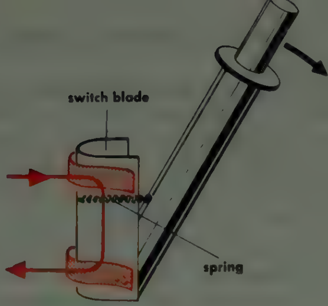


Fig. 2 KNIFE SWITCH

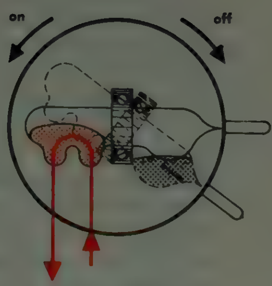


Fig. 3 MERCURY SWITCH

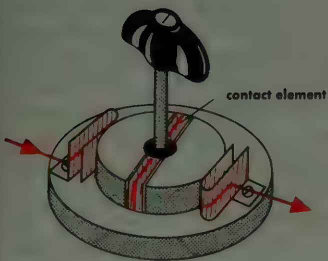


Fig. 4 TURN SWITCH

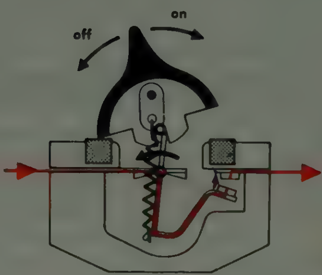


Fig. 5 TUMBLER SWITCH

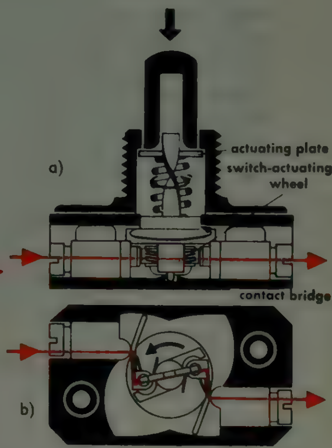


Fig. 6 PUSH-BUTTON SWITCH: *longitudinal section (a), cross-section (b)*

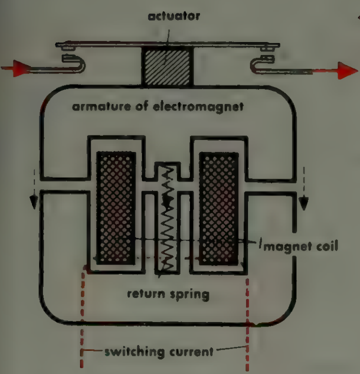


Fig. 7 ELECTROMAGNETIC SWITCH

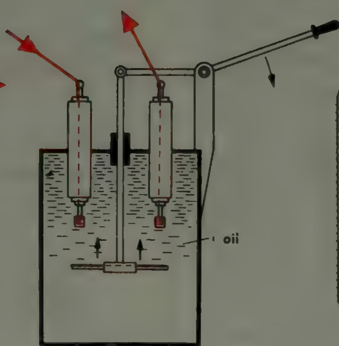


Fig. 8 OIL-BREAK SWITCH

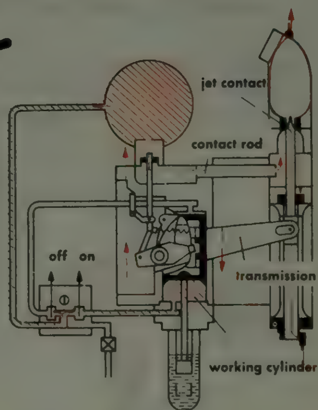


Fig. 9 GAS-BLAST SWITCH

If the poles of an electric current source are interconnected by a short thick metal bar providing an easy path for the current (Fig. 1a), a very strong current will flow and the conductors (wires or cables) will become very hot, with an attendant fire hazard. According to Ohm's Law, the relationship between the voltage (U), the amperage or current strength (I) and the resistance (R) is as follows; $U = I.R$ or $I = U/R$. From this it appears that if the resistance in an electric circuit becomes abnormally low in consequence of some accidental cause, the current strength will increase considerably. This phenomenon is called *short circuit*. It is liable to occur more particularly in faulty leads or in electrical appliances such as a standard lamp whose flexible supply cord ("flex") is subjected to frequent bending, as a result of which the insulation between the two wires becomes damaged and allows them to touch each other, so that short circuit occurs. Damage due to short circuit is prevented by the inclusion of fuses in electric circuits. The simplest kind of fuse is merely a length of thin wire which, in the event of a short circuit, is heated rapidly by the ensuing high current and melts away, thus interrupting the circuit (Fig. 1b). Fig. 2 shows a more sophisticated fuse embodying this principle. The essential component is the fuse cartridge containing the fuse element (fusible wire or strip) embedded in sand. The rating of the fuse element, i.e., the current strength at which it will melt, will depend on the operating conditions and the degree of safeguard to be provided. A "blown" fuse is recognisable by the fact that a small coloured distinctive disc on the front of the fuse has dropped off. The disc serves as a colour code for the fuse rating, e.g., a green disc for a 6-amp. fuse, a red disc for a 10-amp. fuse, etc. When a fuse has blown, the fuse cartridge must be removed and a new cartridge inserted (of course, after the fault which caused the short circuit has been traced and remedied). The need for fitting a new cartridge is obviated in the automatic cut-out (or automatic circuit breaker) (Figs. 3a and 3b). In a device of this kind a switch is held in the "on" position by means of a catch. The catch can be withdrawn either by the action of an electromagnet or a bimetallic strip. When this happens, the switch is released and springs open, thereby interrupting the circuit. The electromagnet is so designed that it is actuated by the current in the circuit as soon as this current exceeds a certain predetermined value. In the other type, the heat produced by the excessive current in the event of short circuit causes the bimetallic strip to bend and thus interrupt the flow of current. When the cause of the short circuit has been eliminated, the automatic cut-out can be re-set in its operating position by actuation of a push-button. A second push button may be provided, to enable the cut-out to be "tripped" manually, so as to break the circuit at will.

Fig. 1 PRINCIPLE OF A FUSE

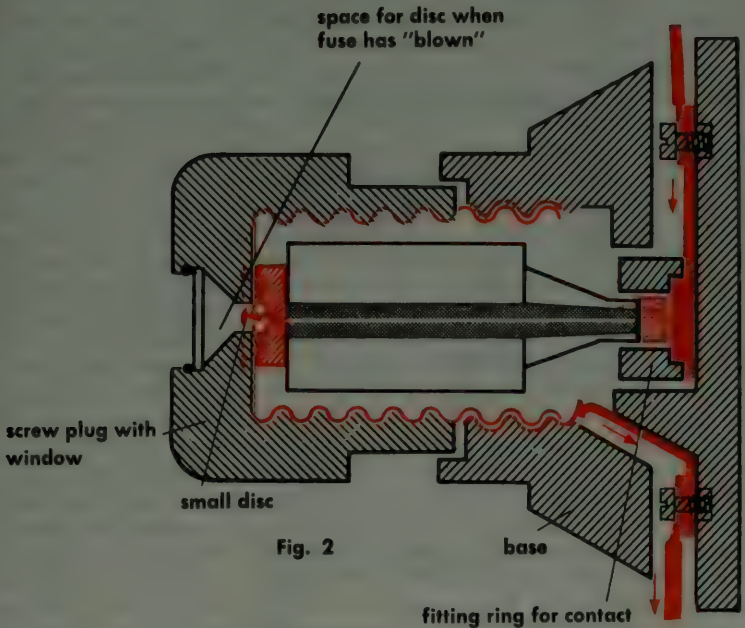


Fig. 2

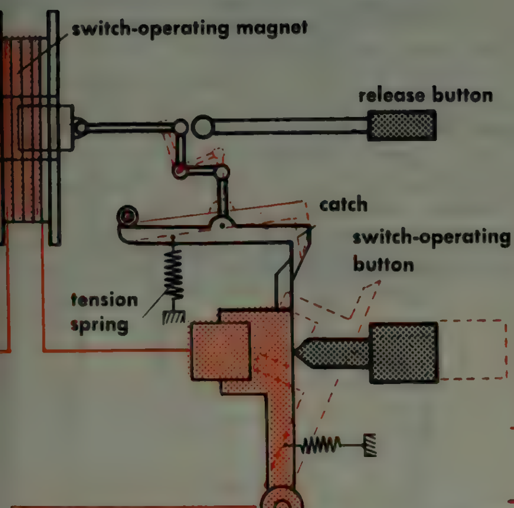


Fig. 3a
AUTOMATIC CIRCUIT-BREAKER
(*electromagnetic type*)

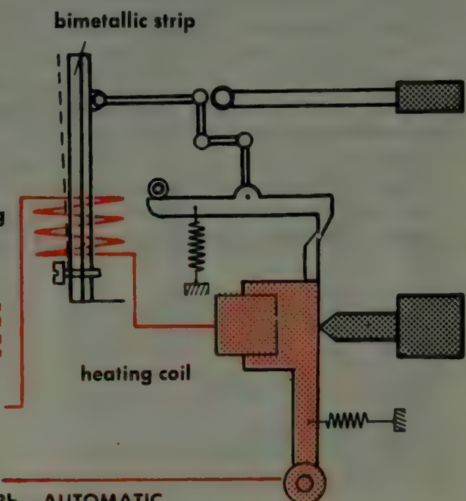


Fig. 3b AUTOMATIC
CIRCUIT-BREAKER
(*thermal type*)

The electric discharge in the form of an arc is allied to the gas discharge which takes place when electricity is passed through rarefied gases and which is the underlying principle of fluorescent lamps (see page 122). The arc discharge occurs when two carbon electrodes are brought into contact with each other and are then moved apart a distance of about one eighth of an inch (the voltage should be at least 55 volts). Just before the carbon rods separate and direct material contact between them is broken, such a high electric resistance is developed at their boundary that the tips of the carbons begin to glow. This is associated with the emission of electrons (see page 90), which, because of the high emission temperatures (up to 4000° C), produces a high degree of ionisation of the air. (With direct current the electrons are emitted from the cathode, i.e., the negative electrode; with alternating current the emission occurs at both electrodes alternately). As a result of this ionisation, the air in the immediate vicinity of the carbon tips becomes conductive to electricity, so that the current will continue to flow when the electrodes are no longer actually touching each other. The bombardment of electrons to which it is exposed causes the positive electrode (anode), in particular, to become white hot, and a "crater" forms at its tip. In the actual arc itself, which merely gives off yellowish violet light, the gas molecules of the air dissociate. They lose some of their enveloping electrons and form a mixture of positive ions (electrically charged atoms) and electrons (negatively charged), which is externally neutral and which, on account of its particular properties, is called *thermal plasma* (Fig. 1). The temperature of this gaseous state can be determined by spectroscopic investigations of its dissociated condition. It is found to be between 20,000° and 50,000° C in the arc. In the arc lamp the arc serves as a source of light, but most of the light comes from the incandescent tips of the carbons (Fig. 2) and especially from the positive crater if the arc lamp is fed with direct current (Fig. 3). As the carbons burn away, they have to be fed forward so as to keep the gap between them fairly constant. If this gap becomes too large, the arc will be extinguished. In modern arc lamps the electrode feed is performed automatically (Fig. 4). The springs F_1 and F_2 keep the carbons in contact with each other when the lamp is not functioning. When the current is switched on, the electromagnets E_1 and E_2 draw the carbons apart and thereby strike the arc. If the rate of burning away is too low, the resistance of the arc will increase. As a result, the current will become weaker, the pull exerted by the electromagnets will diminish, and the springs will draw the carbons closer together. This kind of control mechanism is still sometimes used in arc lamps of cinema projectors, but high-pressure gas discharge lamps are now superseding the arc lamp for this purpose.

In electric furnaces the intense heat developed by the arc discharge is utilised for the melting of metals such as steel. If the material to be melted is a poor conductor of electricity, the heat radiated by the arc formed between two carbon electrodes is used to melt it (Fig. 5). On the other hand, if the material does conduct electricity, then the arc discharge may either be passed direct from the electrodes to the material (Fig. 6) or the electrodes may be actually buried in the material (Fig. 7). In both cases the considerable heat developed in the electrodes helps the current to generate heat in the material and thus attain the melting temperature.

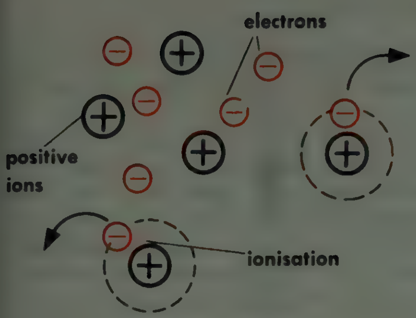


Fig. 1 THERMAL PLASMA

Fig. 2 ALTERNATING CURRENT ARC

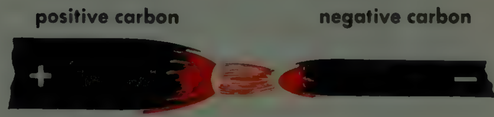


Fig. 3 DIRECT CURRENT ARC

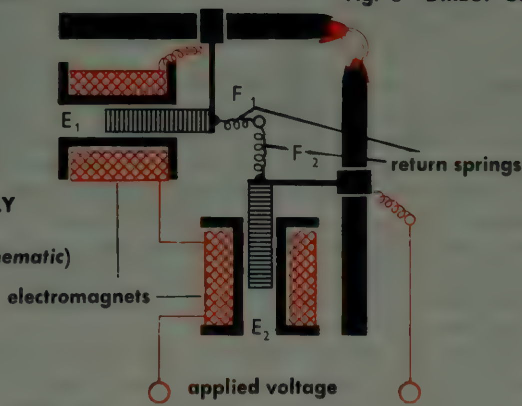


Fig. 4 AUTOMATICALLY CONTROLLED ARC LAMP (*schematic*)

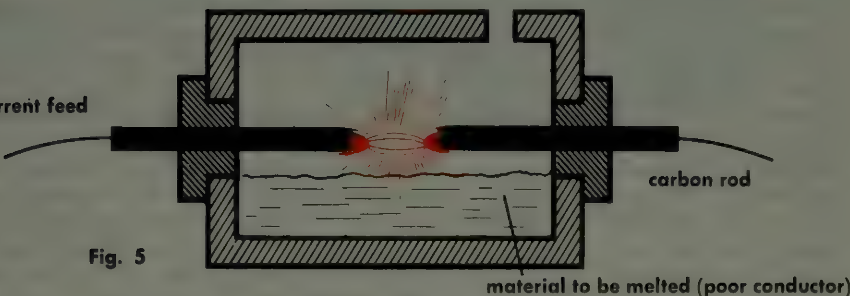


Fig. 5

Fig. 6

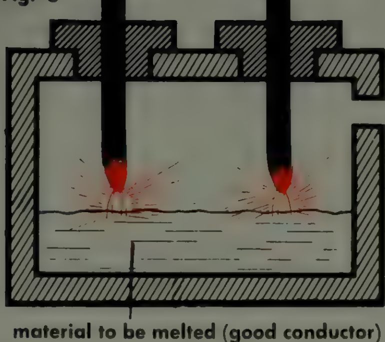
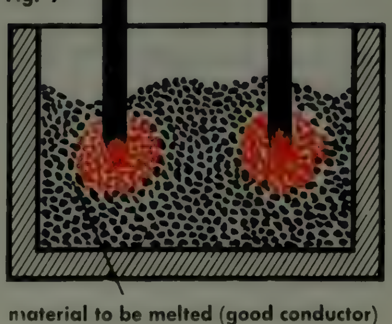


Fig. 7



The functioning of a photo-electric cell (or photocell) which is, for example, an important component of an exposure meter (see page 194), is based on photo-electric effects.

The electrons in a metal can have energy supplied to them by radiation, e.g., light rays. This is known as photo-electric effect. The energy of a light quantum (photon) is imparted to the most loosely bound electron of an atom (Fig. 1a). This energy may be sufficient to liberate the electron but not enough to eject it entirely from the metal (Fig. 1b) (photo conductive effect); alternatively, it may be sufficient not only to liberate the electron but also to cause it to be ejected into the vacuum (Fig. 1c) (normal photo-electric effect). The energy balance of the elementary process involved is given by Einstein's equation:

$$hv = A + \tfrac{1}{2} m_e v^2$$

where hv denotes the energy of the photon, in which v is the frequency of the light radiation and h is Planck's constant ($h = 6.625 \times 10^{-27}$ erg-seconds), A denotes the photo-electric work function (i.e., the energy required by a photon to eject an electron from a metal), m_e denotes the mass of the electron, and v its velocity in vacuum. The normal photo-electric effect is applied in the *photo-electric cell* (Fig. 2a). The light-sensitive photo-cathode, which is usually installed in an evacuated glass tube, may consist of a very thin film of cesium deposited by vaporisation on to an oxidised silver base. For greater sensitivity the glass tube may be filled with an inert gas at low pressure. A battery in the external circuit serves to amplify the current by ionisation of the gas filling.

The photo-conductive effect is utilised in the *photo-conductive cell* (Fig. 2b). The sensitive material usually employed in this case is cadmium sulphide or cadmium selenide. These substances undergo changes in resistance in the ratio of $10^9:1$ between the extremes of darkness and maximum exposure to light.

When the photo-conductive effect occurs at the P-N boundary of semiconductors (see page 98) or at the boundary between a semiconductor and a metal (e.g., cuprous oxide and copper), a potential difference will develop: this is known as the photo voltaic effect, and a cell of this kind is called a photo-voltaic cell (Fig. 2c). The cells represented in Figs. 2a and 2c generate an electromotive force on their own account, causing a current to flow in the circuit even if no battery is included in the circuit, whereas the photo-conductive cell (Fig. 2b) requires an auxiliary voltage provided by a battery. Photo-electric cells are used for a wide variety of purposes in control engineering, for precision measuring devices, in exposure meters used in photography, etc. They are also used in "solar batteries" as sources of electric power for rockets and satellites used in space research. For this purpose silicon photo-electric cells are used; about 10% of the radiation energy which they absorb is converted into electric energy.

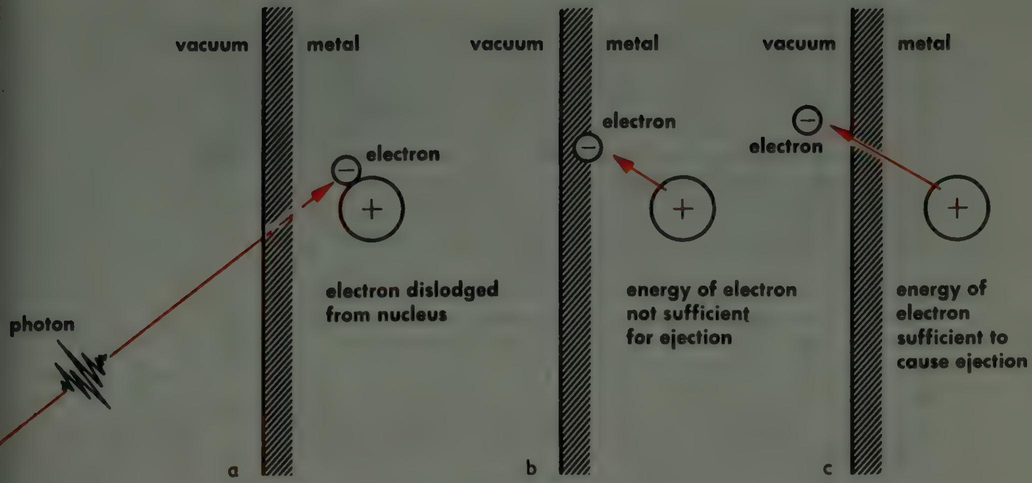


Fig. 1 PHOTO-ELECTRIC PRINCIPLE

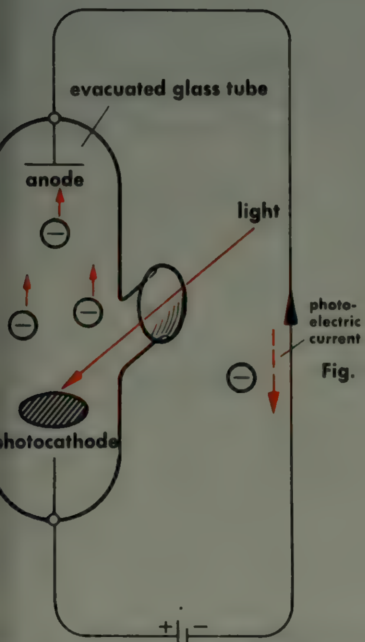


Fig. 2a PHOTO-ELECTRIC EFFECT
(*photo-electric cell*)

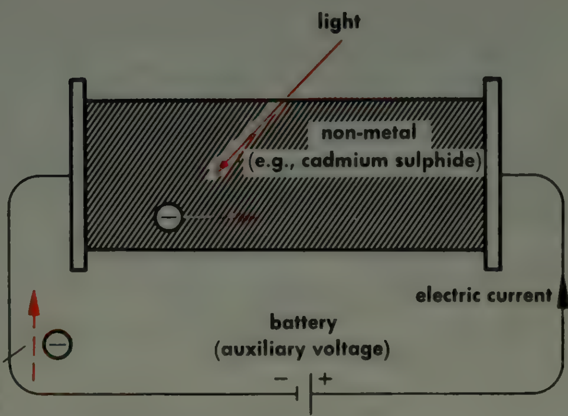


Fig. 2b PHOTO-CONDUCTIVE EFFECT
(*photo-conductive cell*)

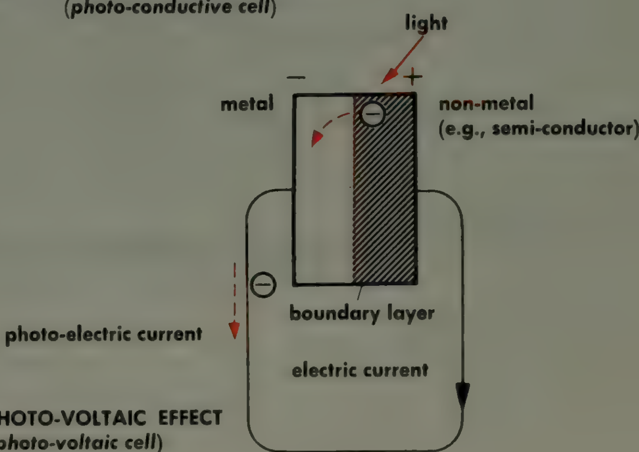
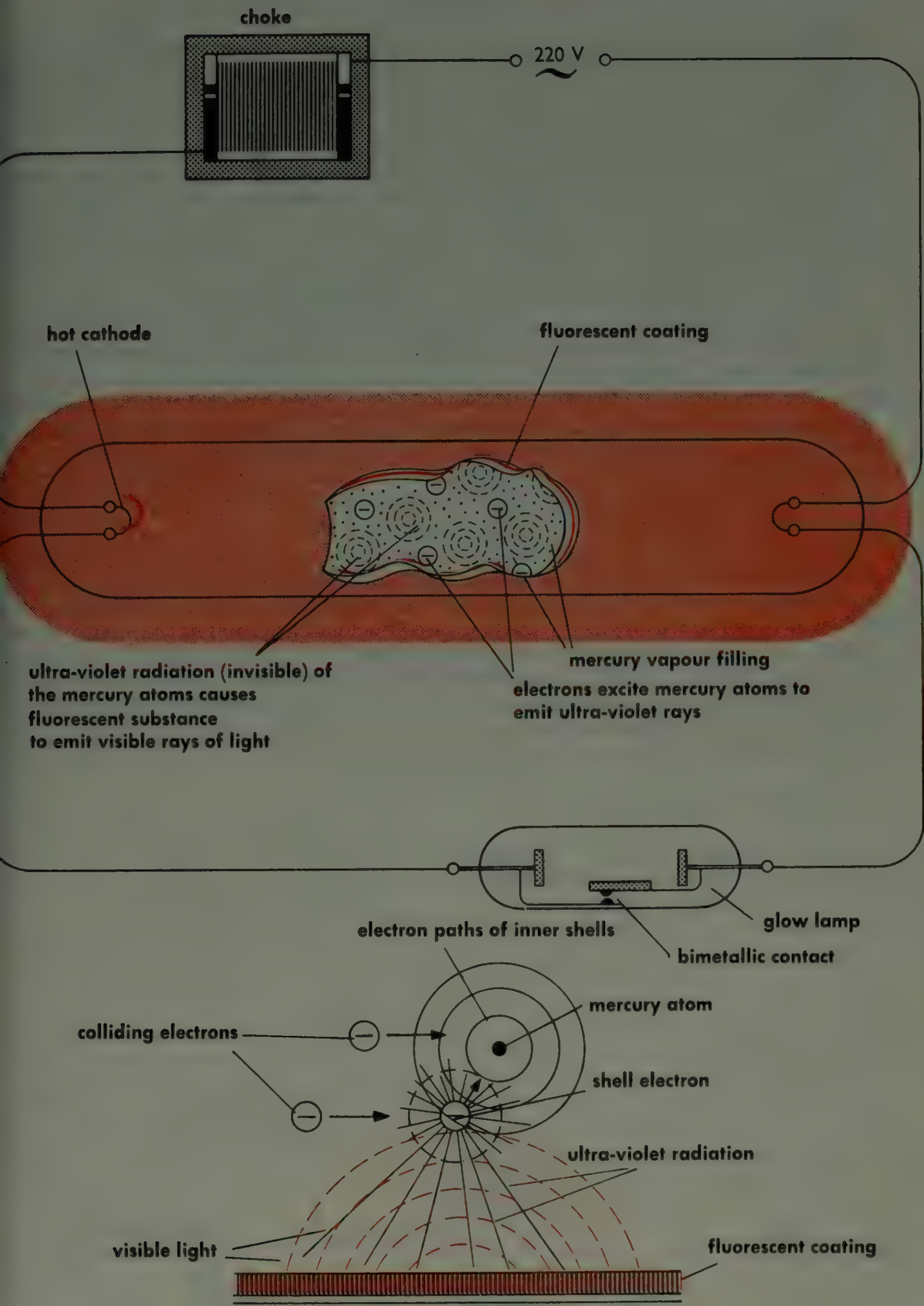


Fig. 2c PHOTO-VOLTAIC EFFECT
(*photo-voltaic cell*)

The fluorescent lamp is a gas discharge tube whose output of light is so increased by special means that it can be used for lighting purposes. The inner surface of the wall of the tube is coated with a light-emitting substance—usually fluorescent or phosphorescent metallic salts (calcium tungstate, zinc sulphide, zinc silicate). The tube is filled with mercury vapour at extremely low pressure. The electrons ejected from the incandescent electrodes (see page 90) collide with the mercury atoms and cause these to emit radiation which consists for the most part of ultraviolet rays, which are invisible. The visible portion of the mercury vapour rays is situated in the green and blue range of the spectrum (see page 156) and gives a pale light. The ultra-violet light strikes the fluorescent substance with which the wall of the tube is coated and causes this substance to emit radiation with a longer wavelength in the visible range of the spectrum—i.e., the coating transforms the invisible rays into visible light. By suitable choice of the fluorescent substance, this light can be given any desired colour. The lamp has to be operated with a choke,[1] which prevents a harmful rise in voltage and serves to ignite the lamp. For this purpose a small auxiliary glow lamp provided with a thermal contact is connected in parallel with the main lamp. When the current is switched on, the glow lamp first lights up (the bimetallic thermal contact is then open). This causes the bimetallic strip to warm up and close the contact, with the result that the glow lamp is short-circuited and the cathodes of the main lamp receive the full current that makes them incandescent. The bimetallic strip cools and breaks the contact. Through the agency of the choke this interruption of the circuit produces a voltage surge which is high enough to initiate the discharge in the fluorescent lamp itself. Because it is bypassed by the main lamp, the small auxiliary lamp then ceases to function. The bimetallic strip (cf. page 26) keeps the contact open. The cathodes of the main lamp are kept glowing at white heat by the impingement of positive mercury ions, and the lamp thus continues to function and emit light in the manner described. The light of a fluorescent lamp is not produced by an incandescent body (such as the filament of an ordinary electric lamp), but is emitted as a result of the excitation of atoms (namely, those of the mercury vapour and the fluorescent coating) and is extremely economical. Because of the large light-emitting area, a fluorescent lamp gives a pleasant light which produces only soft shadows.

1. Starter in U.S.A.


choke
220 V ~
hot cathode
fluorescent coating
ultra-violet radiation (invisible) of the mercury atoms causes fluorescent substance to emit visible rays of light
mercury vapour filling
electrons excite mercury atoms to emit ultra-violet rays
glow lamp
bimetallic contact
electron paths of inner shells
mercury atom
colliding electrons
shell electron
ultra-violet radiation
visible light
fluorescent coating

When a stream of very fast high-energy electrons strikes a metallic electrode (anode), the electrons are slowed down, and some of them penetrate into the metal (Fig. 1). The sudden "braking" of the electrons produces an electromagnetic radiation of very short wavelength: X-rays or Röentgen rays. This radiation is generated by electrons penetrating into the metal and interacting with the metal atoms. It shows well-defined wavelengths which are characteristic of the structure of the metal forming the anode: a high-energy electron which penetrates into the metal atom may dislodge one of the inner electrons of that atom; the vacant place is taken by one of the outer electrons which thus leaps from an outer to an inner "shell" and, in doing so, emits energy in the form of radiation, i.e., X-rays.

These rays were discovered by W. Röntgen, a German physicist, in 1895. Their nature was unknown to him, and he accordingly referred to them as "X-rays", a name which has persisted more particularly in the English-speaking countries. Technical forms of construction of X-ray tubes are illustrated in Figs. 3 and 4. As a rule, the stream of electrons (such electrons issuing from a cathode are called "cathode rays") is not directed against the actual anode, but against the anticathode, which forms a target for bombardment. The impingement of the electrons against the anticathode causes the latter to become very hot, and it may be necessary to cool it or to design it as a rotating anode, so that the cathode rays are always beamed on a fresh area of the anode surface (cf. page 156, vol. II).

Because of their short wavelength (10^{-8} to 10^{-12} cm) X-rays can pass through objects which are opaque to ordinary light, and shadow images of such objects can be made visible on a fluorescent screen coated with barium platinocyanide. When X-rays pass through crystalline substances, diffraction phenomena occur which reveal the wave character of this radiation. An interference pattern (cf. page 156) composed of a regular arrangement of dots can be formed on a photographic plate, and these provide information as to the crystal structure of the material concerned. Such diffraction patterns were first studied scientifically by M. von Laue, and they are known as Laue X-ray patterns (Fig. 5).

Fig. 1

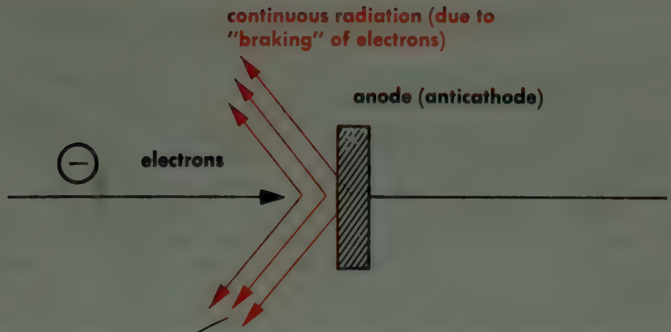


Fig. 2

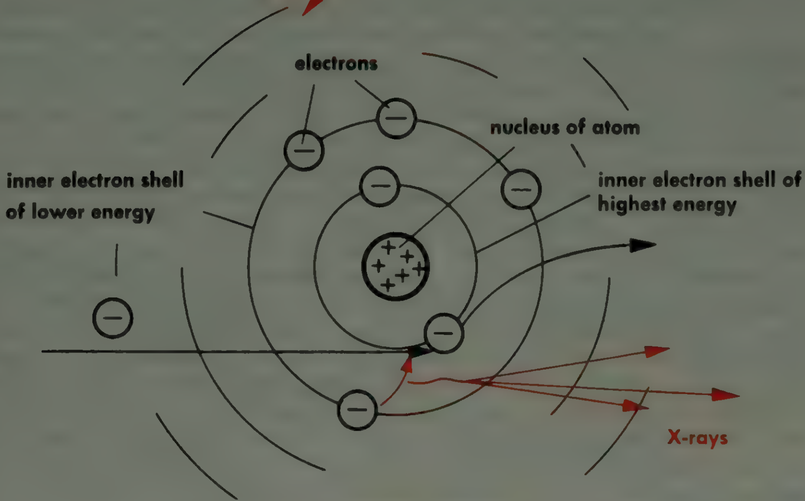


Fig. 3

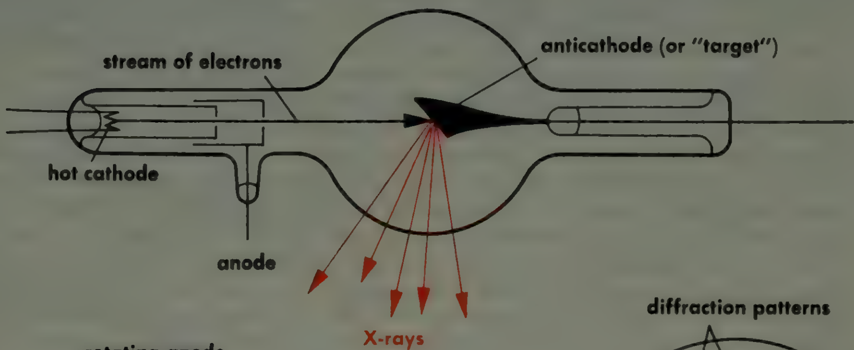
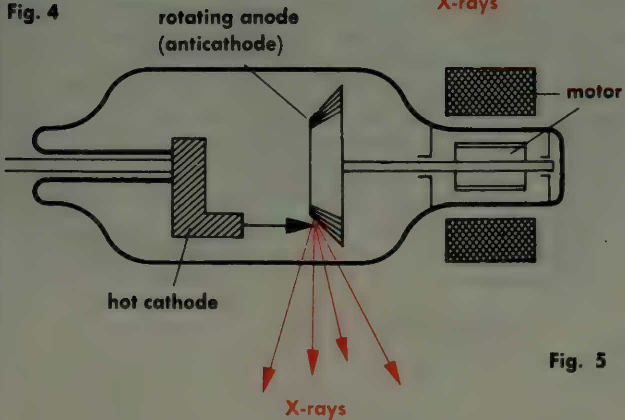


Fig. 4



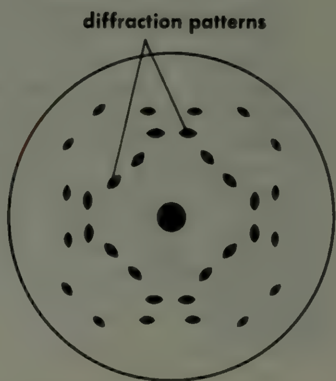


Fig. 5 LAUE X-RAY PATTERN (*zinc blende*)

Scientists engaged in research into the structure of the atom use high-energy "projectiles" for bombarding the atom. Elementary particles (e.g., electrons) make very suitable projectiles because they mostly have an electric charge, so that they can be accelerated by electromagnetic fields. The acceleration is usually produced by causing the particles to travel in circular paths, and for this reason the various machines used for the purpose are known collectively as "circular accelerators".

The *betatron* (Fig. 1) resembles a transformer (see page 110) in its general construction: an iron core is provided with an exciting winding (corresponding to the primary winding of a transformer) through which an alternating current flows. Instead of a secondary winding, however, the betatron has an evacuated annular tube into which electrons are shot. In consequence of magnetic induction (see page 76), an annular alternating electric field (rotational field) is produced in the tube. This field, whose direction is perpendicular to that of the magnetic field, can have a retarding as well as an accelerating effect on the electrons. Since acceleration is required, an intermittent mode of operation is applied. The electrons are shot into the annular tube at the instant when the alternating electric field attains its maximum strength. During the next quarter period of the alternating field strength the electrons are accelerated by the field, while the field strength continuously decreases. The electrons must be removed from the annular tube not later than the end of this quarter period, for otherwise they would be decelerated in the following half period.

According to the conventional theory of electrodynamics the rotational frequency of an electrically charged particle travelling in a circular orbit in a magnetic field is independent of the radius of its orbit. The energy of the particle increases with its velocity. The operation of the *cyclotron* is based on these principles. The evacuated acceleration chamber is mounted in a homogeneous magnetic field between the poles of a powerful electromagnet, as shown in Fig. 2a. The actual vacuum chamber is shown (in plan) in Fig. 2b. Ions (electrically charged atoms) are produced by a source at the centre of the chamber and are forced to travel along circular orbits by the magnetic field. As their circular frequency is constant, the ions can be accelerated by means of a high-frequency alternating electric field of the same frequency. This is done with the aid of two hollow semicircular electrodes (called "dees") to which a high-frequency oscillating voltage is applied. Each time an ion passes from one dee into the other, it is accelerated by the electric field that exists in the gap between the two dees. This does not cause any change in the rotational frequency, but it does increase the velocity of the ion, with the result that the radius of its orbit increases. The ion thus progressively acquires more and more energy and moves into orbits of increasingly large radius, i.e., it moves in a spiral path, until it is finally deflected (by means of a deflecting condenser) out of the vacuum chamber at the periphery.

Whereas the cyclotron is used for the acceleration of particles in the low velocity range, the *synchrotron* (Fig. 3) is an accelerator for particles with velocities approaching the velocity of light. The operating principle of the two machines is very similar, however. As distinct from what happens in the cyclotron, in the synchrotron both the orbit radius of the particle and the rotational frequency remain constant, whereas the magnetic field increases (just as in the betatron). The fact that the energy of the particles increases despite the fact that their velocity remains nearly constant is something that can be explained only with the aid of the relativity theory: when the particle velocity becomes almost equal to the velocity of light, even a small increase in velocity will result in a considerable increase in the mass (and therefore the energy) of the particle. The gain in energy achieved in the synchrotron is therefore due, not to a velocity increase, but to an increase in mass.

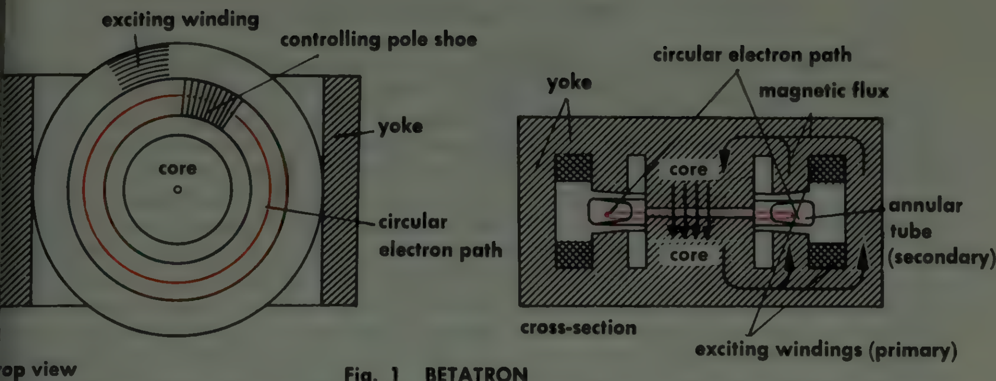


Fig. 1 BETATRON

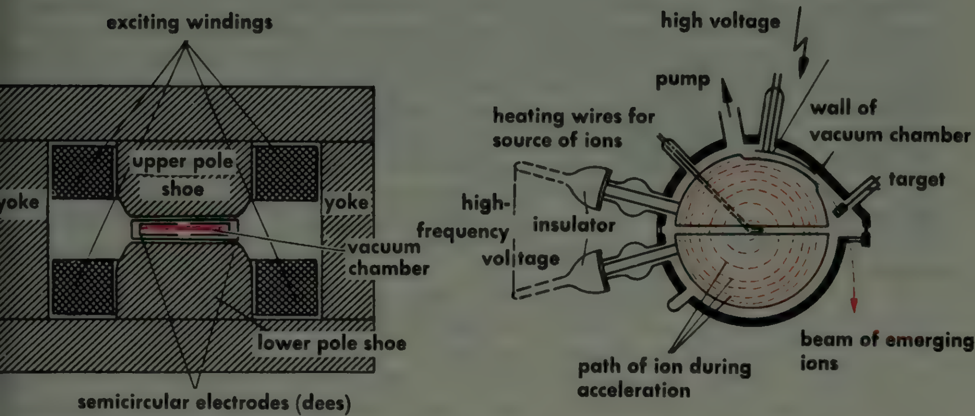


Fig. 2 CYCLOTRON

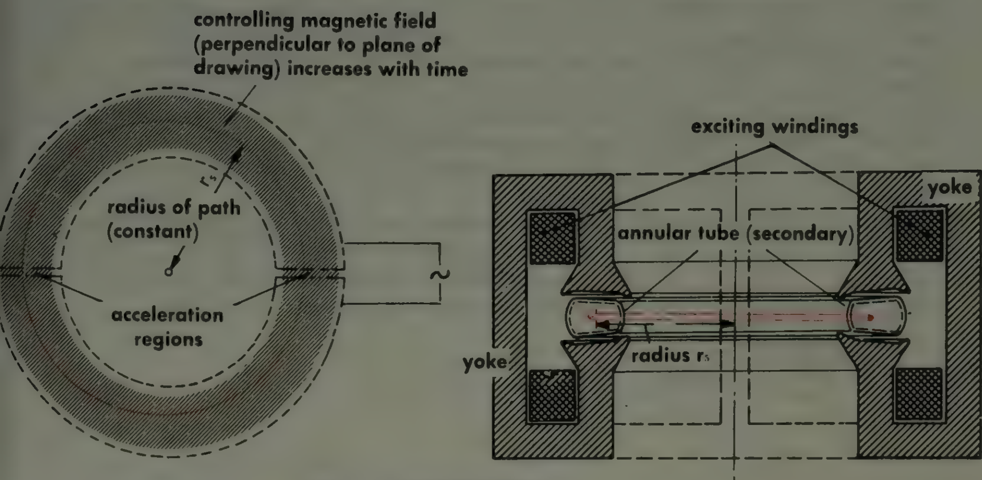


Fig. 3 SYNCHROTRON

Telecommunication, in its widest sense, is as old as mankind's need for ways and means of exchanging information over long distances. In ancient times and during the middle ages, visual signals were usually employed for the purpose, in the form of beacon fires lit on top of towers or on hills and visible a long way off. Of course, only a limited number of prearranged messages could be communicated by such means. A later and more sophisticated method of visual signalling was the semaphore, a device provided with movable arms which could be set in various positions to spell out letters and words. Signalling by means of flags or lamps, and the hoisting and lowering of signal balls, as are still commonly employed in navigation, also belong to the general category of communication by visual signals. Telecommunication in the modern sense developed only after the discovery of the magnetic effect of electric current. The first telegraph consisted of a compass needle which was deflected by the magnetic field produced by electric currents which flowed through the circuit whenever the transmitting key was depressed and contact established (Fig. 1). A further advance in electrical telecommunication was the invention of the Morse telegraph: an electromagnetically actuated stylus records long and short dashes—forming a code named Morse code after its inventor—on a moving strip of paper (cf. page 134). In the latter half of the last century a further advance was made by the invention of the telephone (and microphone) so that it then became possible to transmit speech-modulated current fluctuations over long distances (Fig. 3).

The current fluctuations are transmitted either through overhead wires, through cables, or through the medium of radio communication (Fig. 4). The electric current or the radio waves can be modulated in various ways to carry the message: amplitude modulation (Fig. 5), i.e., variations in current strength, or frequency modulation, i.e., variations in the timing of the zero values of the current (Fig. 6), or pulse modulation (Fig. 7).

As the transmission of messages over long distances, whether by cable or by radio, is very expensive, the problem of simultaneous multiple utilisation had to be tackled. It was solved by means of carrier wave communication: a high-frequency current is modulated in various frequency ranges. For telephony a band width of 3600 cycles/sec. is adopted for each range and is adequate for the intelligible transmission of speech. Each range of this kind is comparable to a wire or cable and is called a channel. The frequency band of a telephony channel can, for example, be subdivided into 24 telegraphy channels. In this way it is possible to cater for the needs of subscribers to the teleprinter (telex) system. Filtering out the individual channels is done by means of electric filter systems. For the transmission of the high-frequency carrier oscillations the cables are of the so-called coaxial type (Fig. 8): a flexible metal sheath contains a coaxial conductor which is held in position by highly insulating material. As a rule, the cable also contains various subsidiary wires for low-frequency, i.e., direct, transmission purposes. In a coaxial cable of this kind up to 2880 telephony channels can be accommodated in the coaxial conductor and in the subsidiary wires. The coaxial core (tube) is additionally suitable for the transmission of a television channel.

Fresh possibilities in electric telecommunication have been opened up by pulse techniques and other modern developments which form the basis of data processing (see page 334). Also, practical applications of masers and lasers (see page 94) are likely to leave their mark on present-day telecommunication methods.

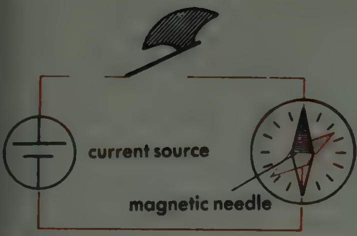


Fig. 1

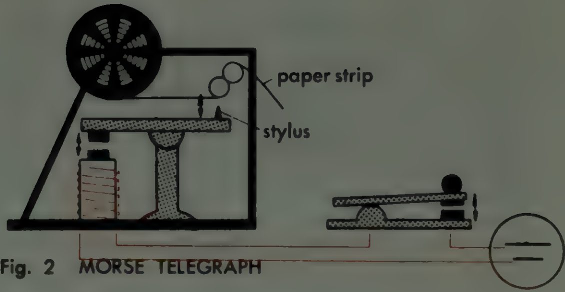


Fig. 2 MORSE TELEGRAPH

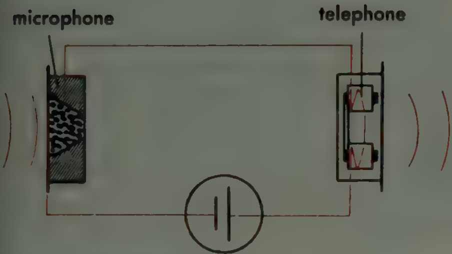


Fig. 3 TELEPHONY

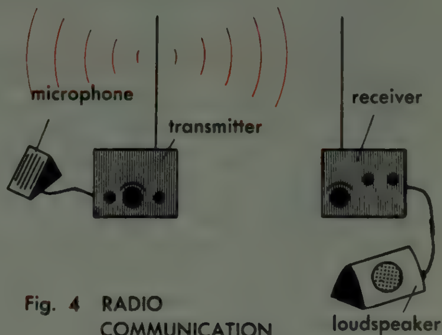


Fig. 4 RADIO COMMUNICATION

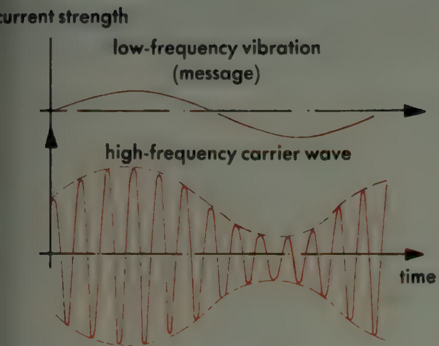


Fig. 5

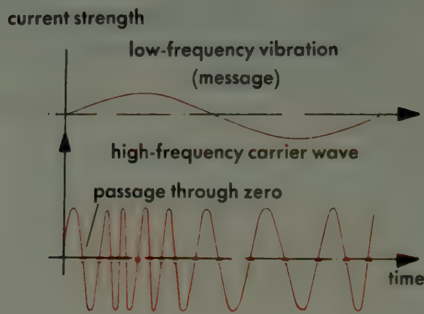


Fig. 6

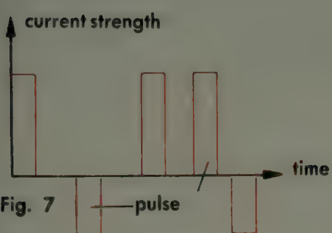


Fig. 7

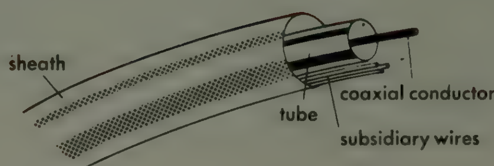


Fig. 8 COAXIAL CABLE

The basic problem in telephony consists in converting acoustic into electric energy and vice versa. The first part of the problem is solved by means of the carbon microphone in which carbon granules are compressed to a greater or less degree by a diaphragm, so that their resistance varies with the acoustic pressure. This in turn produces corresponding fluctuations in the current, which are transmitted—through a transformer called a repeating coil—to the telephone receiver. The latter comprises a U-shaped permanent magnet whose legs are provided with coils which are energised by the telephone current. The fluctuations in the strength of the magnetic field, and these correspondingly move the steel diaphragm, which in turn transmits its vibrations to the air in the form of audible sound waves (Fig. 1). As an alternative to the carbon microphone the condenser microphone may be employed (Fig. 2). This is a more sensitive instrument. Its diaphragm, in conjunction with a fixed counter electrode, forms a condenser whose capacity varies with the vibrations of the diaphragm. The direct voltage applied in the circuit thereby has an alternating voltage superimposed upon it.

The interconnection of two subscribers in the central battery system is indicated diagrammatically in Fig. 3. With manual switching (Fig. 4) the connection with the desired subscriber is effected through various exchanges (local exchange, trunk exchange, zone centre, trunk exchange, local exchange). The connections are routed through subscriber's lines (*AL*), interoffice trunks (*VL*), and trunk lines (*FL*). In the automatic long distance service the connections are nowadays still usually established by means of mechanical selector units (two-motion selectors). When the subscriber dials the desired number, electrical impulses, corresponding to the respective numbers, are transmitted along the telephone line and actuate a selector at the telephone exchange (Figs. 5 and 6).

A selector is an automatically actuated multiple contact switch or other device which makes contact with any desired circuit. In the two-motion type the wiper can move vertically step by step from one tier of contacts to another, and at each tier it can swing horizontally to any particular contact. These movements are controlled by the current impulses transmitted by dialling. For each figure dialled, a corresponding number of impulses of current reaches the selector, causing the wiper to move automatically to the correct tier and to the correct individual contact corresponding to the figure dialled.

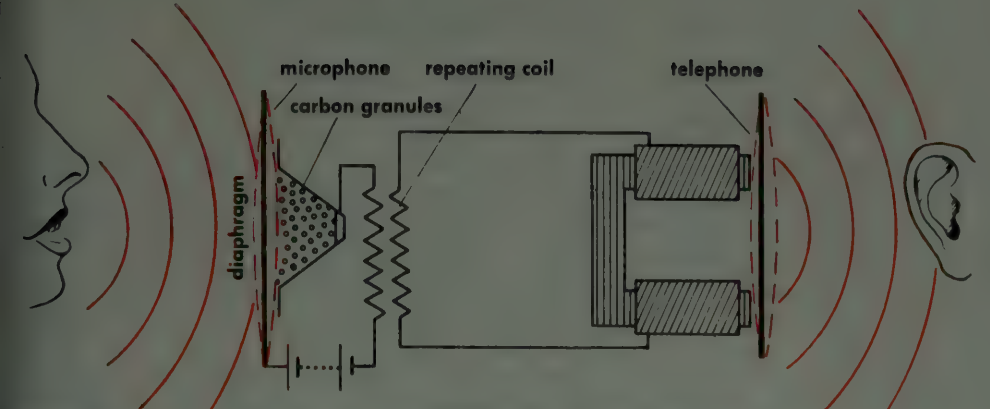
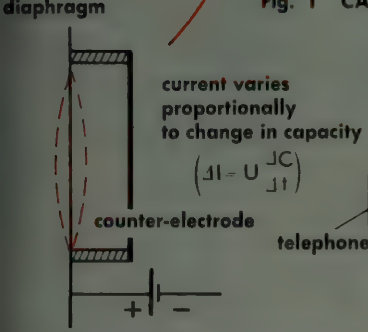


Fig. 1 CARBON MICROPHONE AND TELEPHONE



g. 2 CONDENSER MICROPHONE

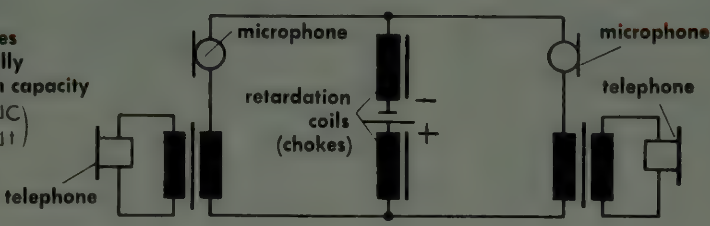


Fig. 3 CENTRAL BATTERY SYSTEM

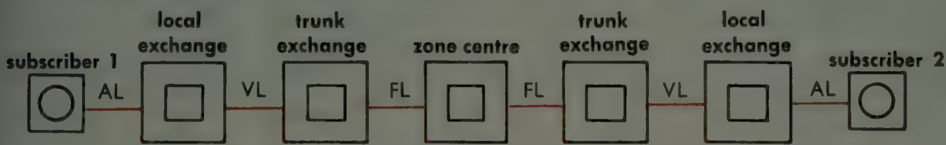


Fig. 4 MANUAL SWITCHING

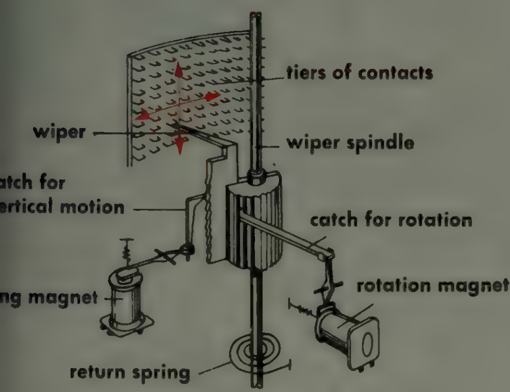


Fig. 5 TWO-MOTION SELECTOR

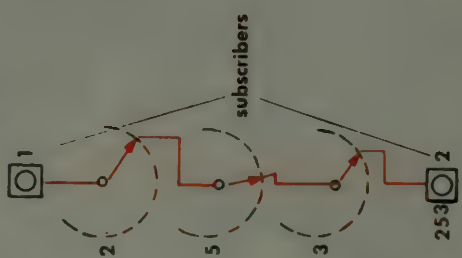
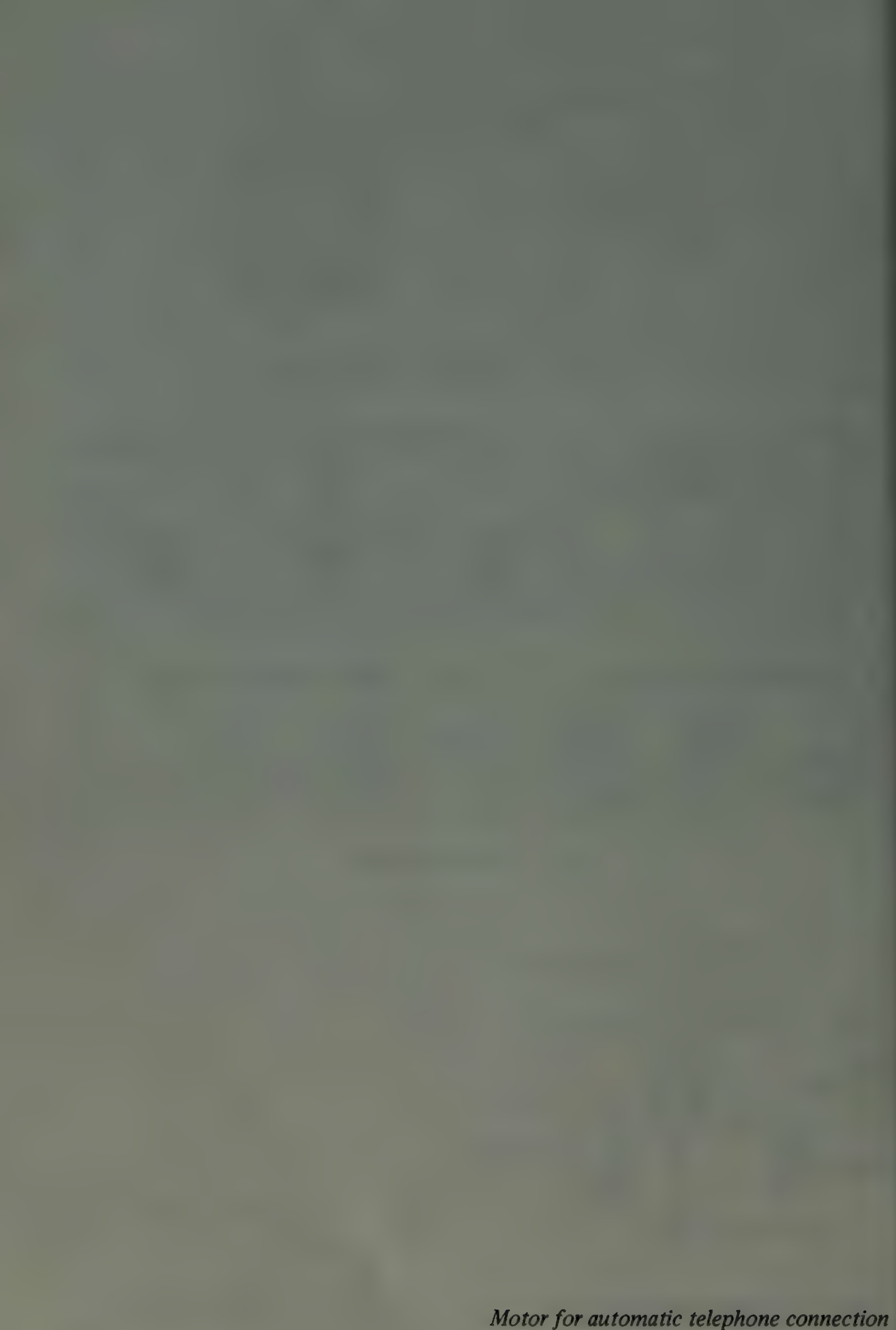
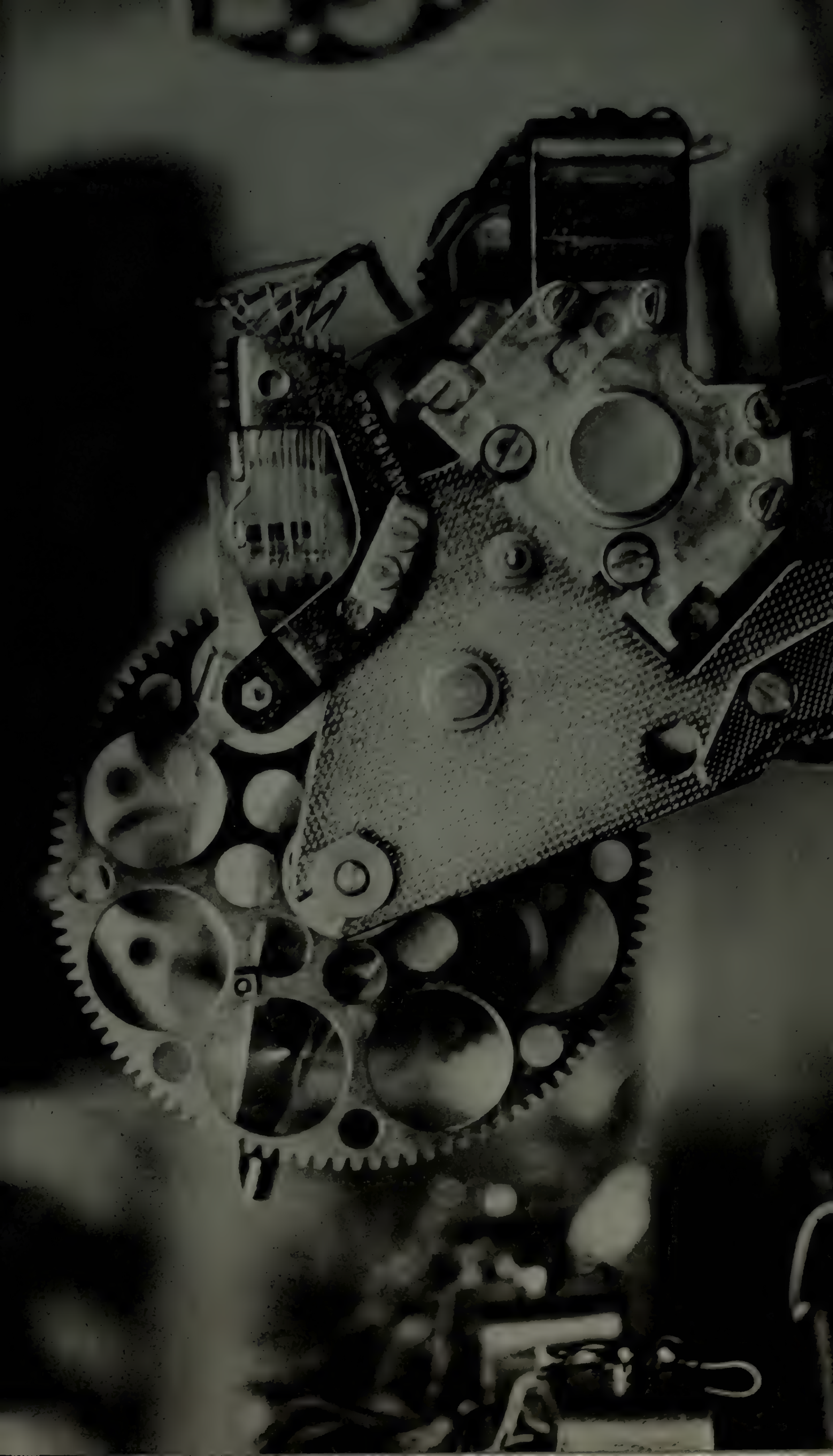


Fig. 6 SELECTOR UNIT FOR
THREE-STAGE CONNECTION



Motor for automatic telephone connection
Photo Roland Schneider, Len Sirman Press



After the discovery of electromagnetism in the first part of the nineteenth century, inventors strove to utilise it for the transmission of messages by electricity. After the promising tests carried out by Gauss and Weber at the University of Göttingen, Germany, in 1833, Morse developed a reliably functioning telegraph apparatus and a workable code (Morse alphabet). In its present form the Morse apparatus has a strip of paper which is unwound from a supply reel and slides past an inked printing wheel. When at rest, this wheel does not touch the paper strip, but when a message is being received, the wheel is pressed against the moving strip for shorter or longer periods of time by means of an electromagnet, so that a series of "dots" and "dashes" are produced (Fig. 1). The International Morse code is as follows:

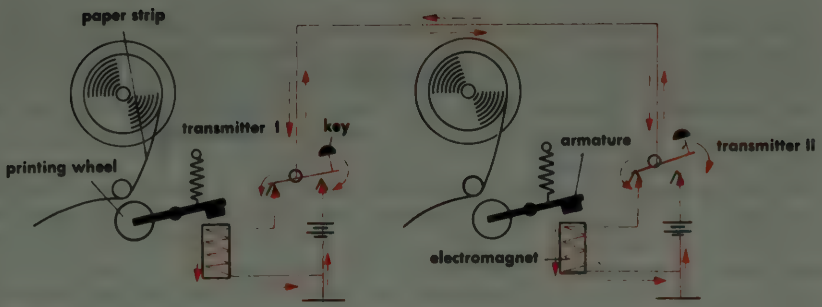
Alphabet						*Numerals*	
a	· –	h	· · · ·	q	– – · –	1	· – – – –
ä	· – · –	i	· ·	r	· – ·	2	· · – – –
å	· – – · –	j	· – – –	s	· · ·	3	· · · – –
b	– · · ·	k	– · –	t	–	4	· · · · –
c	– · – ·	l	· – · ·	u	· · –	5	· · · · ·
ch	– – – –	m	– –	ü	· · – –	6	– · · · ·
d	– · ·	n	– ·	v	· · · –	7	– – · · ·
e	·	ñ	– – · – –	w	· – –	8	– – – · ·
é	· · – · ·	o	– – –	x	– · · –	9	– – – – ·
f	· · – ·	ö	– – – ·	y	– · – –	0	– – – – –
g	– – ·	p	· – – ·	z	– – · ·		

Punctuation marks					
apostrophe	· – – – – ·	hyphen	– · · · · –	question mark	· · – – · ·
colon	– – – · · ·	parenthesis	– · – – · –	quotation mark	· – · · – ·
comma	– – · · – –	period	· – · – · –	*Distress signal* (*SOS*)	· · · – – – · · ·

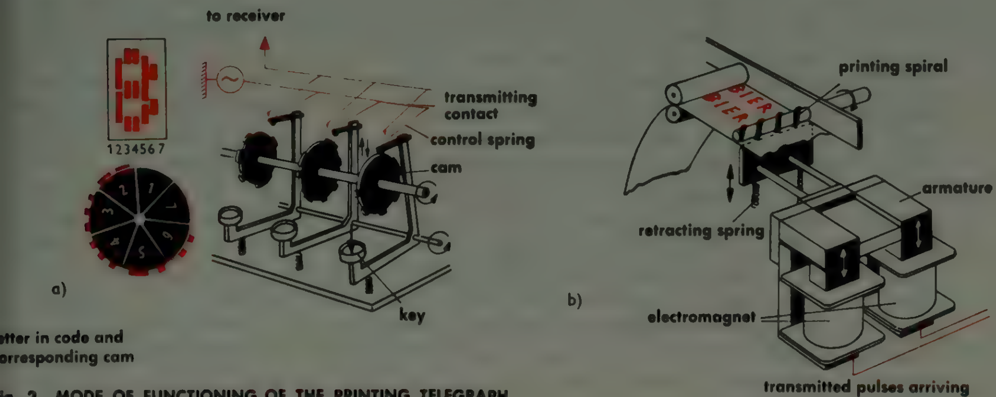
Subsequent inventors aimed at achieving direct transmission of characters, i.e., letters of the alphabet, instead of having to use a code. The latest stages in this process of development are the Hell printing telegraph (Fig. 2) and the teletypewriter (or teleprinter) (Fig. 3). Both these devices make use of an alphabet produced by current impulses on the step-by-step principle. This alphabet operates with a five-unit code comprising seven current steps and five separating steps. The trains of impulses are timed by means of synchronised camshafts rotating in the transmitting and in the receiving instrument respectively (7 r.p.m.). The various code signals (characters) correspond to different notches in the permutation bars in the transmitter. When the keys are depressed, these bars shift to different positions and thereby control the current impulses sent (Fig. 3a). In the receiver, slotted combination bars are similarly so displaced in relation to one another that only that type bar which corresponds to the transmitted code signal engages with the slots in the bars and thus prints a particular character on a sheet or strip of paper (Fig. 3b). These bar movements are produced by an electromagnet. The teletypewriter looks like a large desk typewriter. One and the same apparatus can be used for transmitting and receiving.

Another branch of telegraphy is picture telegraphy (or phototelegraphy[1]): a photograph attached to a revolving drum is scanned point by point in a spiral pattern by a beam of light. The reflected light causes photo-electric cells (see page 120) to produce currents of varying strength which correspond to the variations in brightness of the individual points of the photograph (Fig. 4a). At the receiving end, a glow lamp transforms the current variations into variations of brightness of a light spot (P) which is focused on to photographic paper (Fig. 4b).

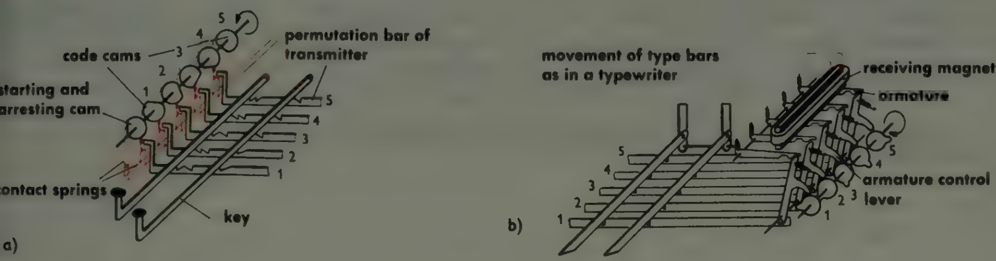
1. Wire-photo in U.S.A.



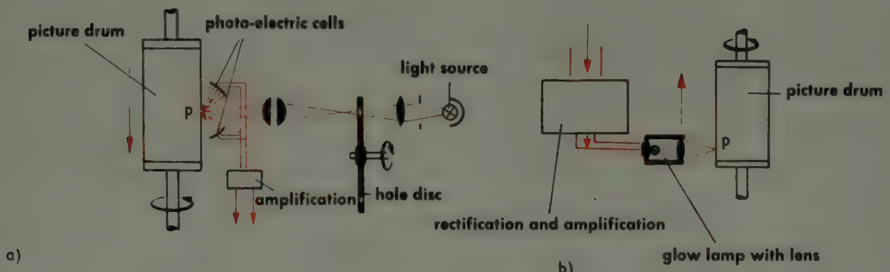
g. 1 MORSE TELEGRAPHY WITH OPEN-CIRCUIT WORKING



ig. 2 MODE OF FUNCTIONING OF THE PRINTING TELEGRAPH



ig. 3 MODE OF FUNCTIONING OF THE TELETYPEWRITER



ig. 4 MODE OF FUNCTIONING OF THE PHOTOTELEGRAPH

The name "radar" has been derived from the initial letters of the phrase "radio detecting and ranging'. It denotes a method of scanning the surrounding space by means of high-frequency radio waves which are sent out from a powerful transmitter and are reflected by any objects which they encounter. The reflected beam is picked up by a receiver; its strength and direction gives information on the size, distance, altitude, etc. of the object.

If, for example, an observer in an aircraft wishes to survey by radar the terrain over which he is flying (Fig. 1), a rotating radar beam is directed downward from the aircraft. The beam scans a circular area in the form of a sector which sweeps round and round. Depending on the nature of the reflecting objects (in this case these are located on the surface of the earth), the intensity of the reflected beam will vary (Fig. 2). The transmission and reception of the high-frequency waves are effected in the radar apparatus (Fig. 3). The radar waves are generated in the transmitter, which is equipped with radio tubes of special design (klystron, magnetron: see page 92 *et seq*). The transmitting antenna usually also functions as the receiving antenna (periodic change-over). The reflected beam is picked up by the receiver and the corresponding electric currents are used to deflect an electron beam in a cathode-ray tube (see page 148). The beam is so deflected that it scans the luminescent screen from the centre to the edge while it rotates at the same speed as the antenna. An echo picked up by the receiver strengthens the flow of electrons in the tube, causing a point of light to appear on the screen and to remain visible by phosphorescent afterglow until fresh echoes are picked up on the next revolution of the scanning antenna. In this way the points of light build up a picture of the area (or space) scanned by the radar beam. The brightness of the display of the signal (the radar echo) on the luminescent screen of the cathode ray tube depends on the reflecting power of the object with regard to the high-frequency radio waves sent out by the radar transmitter. For this reason a radar image generally looks quite different from an optical image, though as a rule they will have the same outlines (Fig. 4).

Most radar sets employ pulse radar. This is so called because the transmitter sends out short intense bursts or pulses of energy with a relatively long interval between pulses. The receiver is active during this interval. When sufficient time has elapsed to permit the reception of echoes from the most distant objects of interest, the transmitter sends another short pulse, and the cycle repeats.

Fig. 1

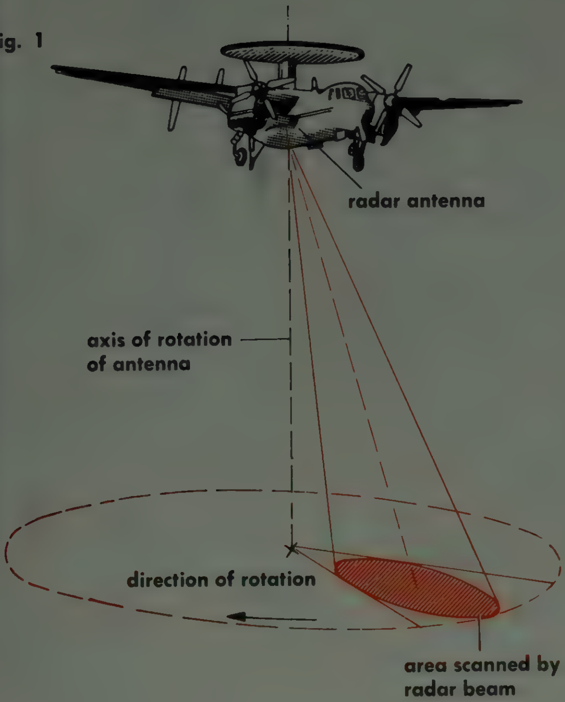


Fig. 2

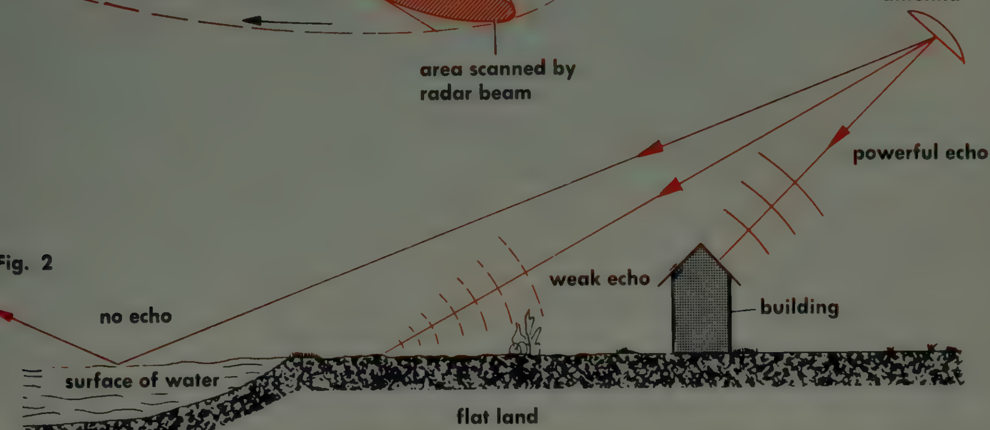


Fig. 3 DIAGRAM OF A RADAR INSTALLATION

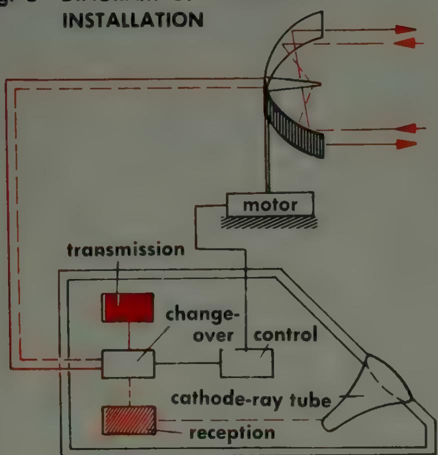
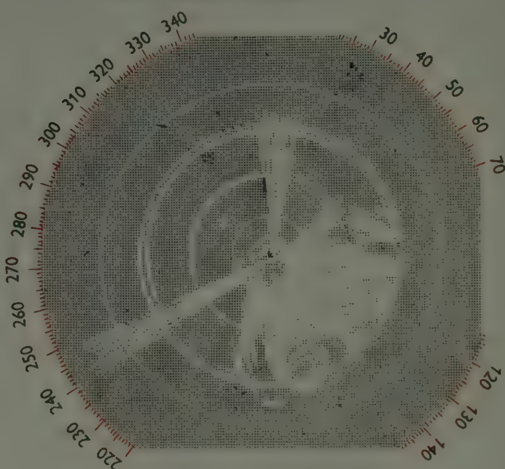


Fig. 4 HOW A HURRICANE APPEARS ON A RADAR SCREEN



Radar antennae
Photo Pierre Berger, Len Sirman Press

An electron multiplier is a device which employs secondary emission (the ejection of electrons as a result of the impact of charged particles) from solids to produce current amplification. A photomultiplier is an electron multiplier tube in which the bombarding electrons initiating the cascade are due to photo-emission, i.e., the emission of electrons from a substance by subjecting it to electro-magnetic radiation such as light, X-rays, etc. Magnesium oxide and caesium oxide layers exhibit a particularly strong secondary emission of electrons. The number of secondary electrons released depends upon the kinetic energy of the primary electrons (the bombarding electrons) and therefore upon the voltage whereby the primary electrons are accelerated. At low voltages, on an average, less than one secondary electron is released per incident primary electron. Only when the secondary emission factor is larger than unity, i.e., when the number of secondary electrons is greater than the number of primary electrons, can one speak of "multiplication". This usually occurs at acceleration voltages above 100 volts (for the primary electrons). For magnesium and caesium oxide layers the secondary emission factor has values above 10, i.e., a more than tenfold multiplication is obtained. The multiplication can be further increased by connecting a number of secondary emission electrodes in series.

If the beam of photo-electrically liberated primary electrons is allowed to impinge upon a row of electrodes provided with a secondary emission coating, amplification values up to 10^9 times the primary radiation can be obtained. The irregularity of the electron flow which is affected by thermal phenomena (thermal noise) sets a limit to further multiplication. The electrodes can be formed as wire gauze electrodes (Fig. 1) or as hollow electrodes (Fig. 2). If need be magnetic deflection can be employed (Fig. 3). In nuclear physics the photo-multiplier plays an important part in the recording of scintillations which are produced by high-energy particles. It is also used as a highly sensitive photometer. In television engineering the photo-multiplier is used in connection with the transmission of cine films and as an important component of camera tubes based on the orthicon principle (see page 146).

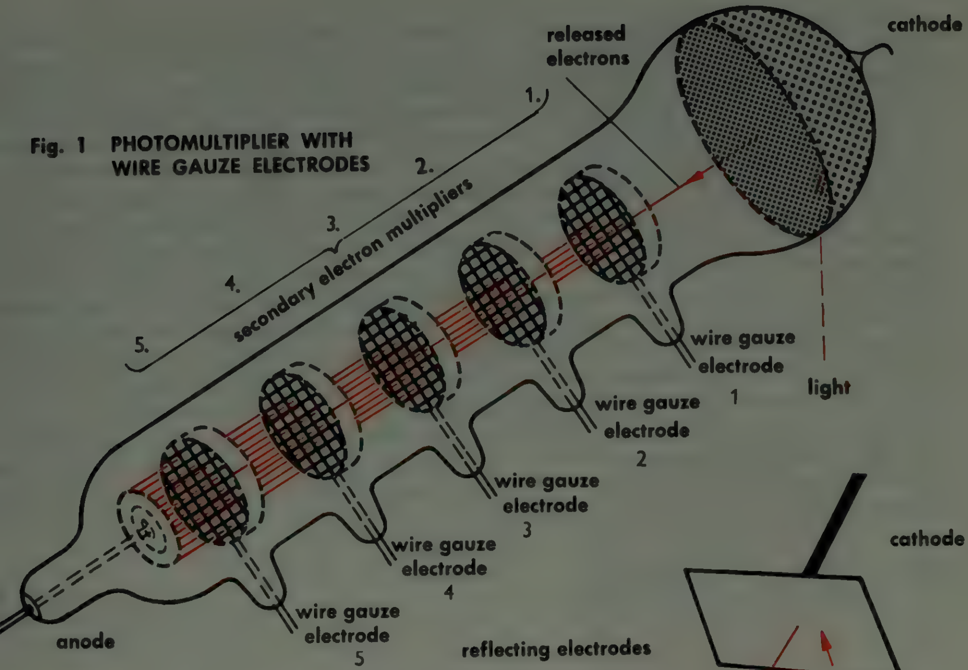


Fig. 1 PHOTOMULTIPLIER WITH WIRE GAUZE ELECTRODES

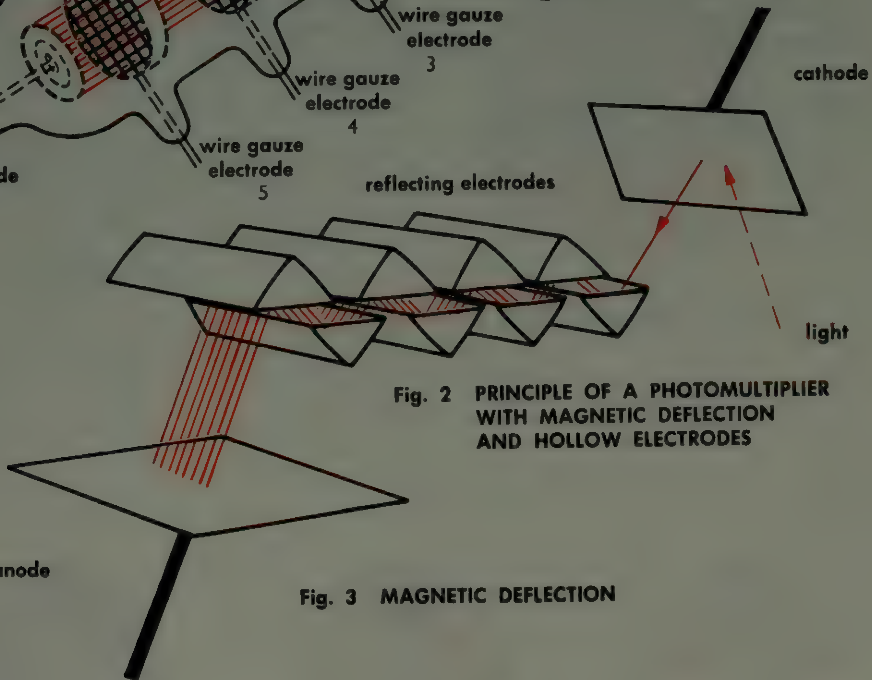
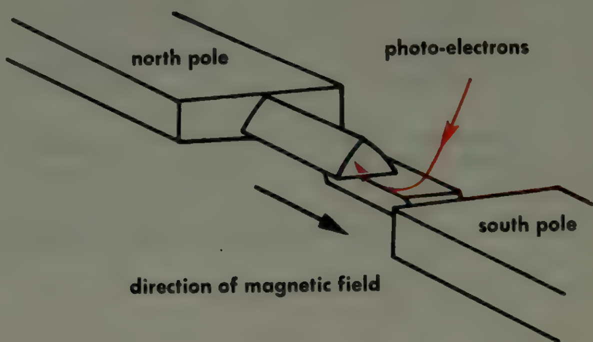


Fig. 2 PRINCIPLE OF A PHOTOMULTIPLIER WITH MAGNETIC DEFLECTION AND HOLLOW ELECTRODES

Fig. 3 MAGNETIC DEFLECTION



The name "magic eye" is popularly applied to the electron-ray tube, a device which gives a visual indication of correct tuning in radio receivers and correct adjustment of the microphone output to obtain a good recording in tape recorders. The electron-ray tube consists of a triode (an electron tube with cathode, control grid and anode: see page 90) and a cathode ray tube (see page 148). The latter comprises a fluorescent screen which is bombarded by electrons emitted by an incandescent cathode; they are accelerated by an anode and are controlled by another electrode (indicator grid). The two parts, i.e., the triode and the cathode ray tube, have separate grids, but share a common, indirectly-heated cathode and have a common anode voltage. The anode of the electron-ray tube comprises two control fins which are conductively connected to the anode of the triode. The fluorescent screen symmetrically surrounds the dark red glowing indirectly-heated cathode, whose light is screened by a cap. The image on the screen in the zero position is shown in Fig. 1b: wide dark sectors are separated by narrow luminous ones. When the receiver is correctly tuned, the luminous sectors open to maximum width.

A further development of this device is the dual electron-ray tube, which allows of coarse and fine adjustment (for powerful and medium-power transmitters). It contains two triode systems and two pairs of anode fins. The fluorescent pattern for various conditions is illustrated in Figs. 2a, 2b and 2c.

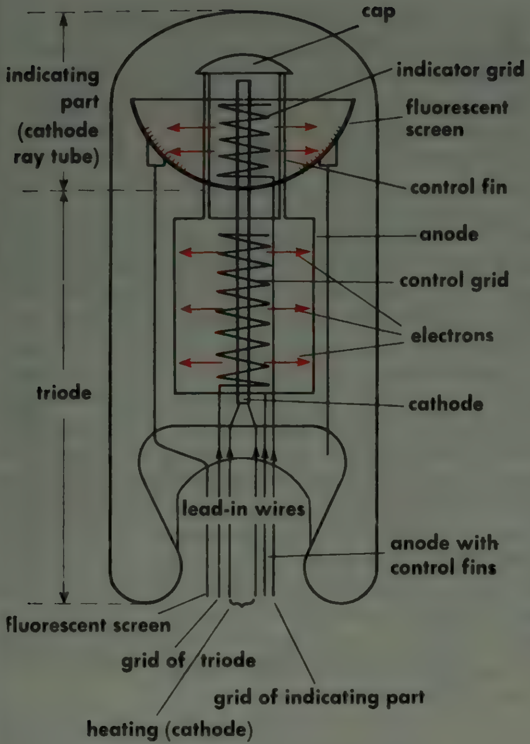


Fig. 1a MAGIC EYE (*schematic*)

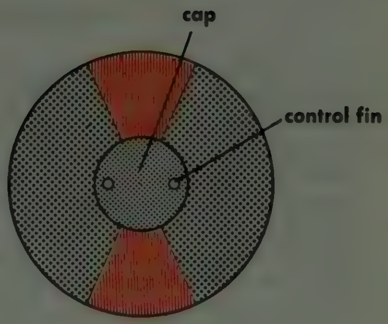


Fig. 1b ZERO POSITION

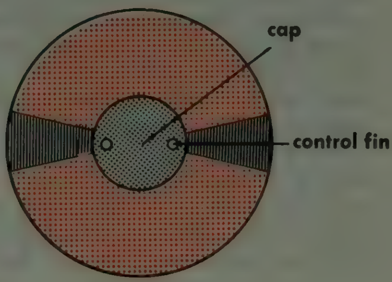
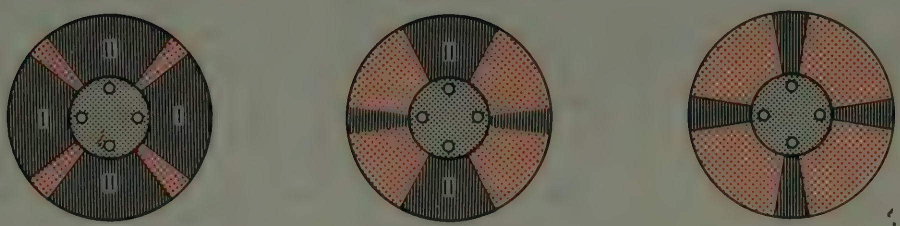


Fig. 1c CORRECTLY TUNED



a) zero position

b) range I adjustment (medium-power transmitter)

c) range II adjustment (powerful transmitter)

Fig. 2 IMAGES FORMED BY DUAL ELECTRON-RAY TUBE

Television utilises the cinematographic projection principle in that it, too, operates at a picture rate of at least 25 per second and thus produces the visual impression of continuous motion. Just as in the half-tone screens used in photo-engraving (see p.128,II), in which the different light and shade values of a photograph are reproduced by a pattern of dots, in television the image is likewise analysed into a large number of "picture elements", the principle of which is illustrated in Fig. 1. This means that the picture must be divided into a number of lines (e.g., 625), and each line must contain several hundred individually identifiable half-tone light values. This is known as "scanning". To obtain a reasonably good picture, the image must be thus analysed into at least 100,000 (and preferably 200,000) picture elements. In the television camera (iconoscope, Fig. 2; for further details see page 146) the image is focused on to a plate called the signal plate whose surface is covered with a mosaic of photosensitive points. Each of these points, corresponding to one picture element, acquires a positive electric charge whose magnitude depends on the strength of the illumination falling on it (Fig. 3). An electron beam, forming a scanning spot on the signal plate, zig-zags its way, line by line, across the plate every 1/25 second[1] and thus discharges each photosensitive point 25 times per second. Each point thus gives an electric impulse whose strength corresponds to the strength of illumination at that point at that particular instant. These impulses (forming the picture signal) are amplified and transmitted. In the television receiver the incoming impulses, after amplification, are fed to the control electrode of the picture tube (cathode ray tube; see page 148)(Fig. 4) in which an electron beam is zig-zagged across a fluorescent screen synchronously with the beam in camera tube and with an intensity varying with the strength of the electric impulses. In this way a pattern of luminous points of varying brightness, and formed in rapid succession, is produced on the screen, thus making the picture that the viewer sees.

The picture signal (Fig. 5) can be conveyed to the receiver by cable (coaxial cables are employed), but they are usually transmitted by means of waves similar to those used in ordinary radio broadcasting, but of shorter wavelength. These high-frequency short waves are only able to travel in straight paths from the transmitter, so that, because of the earth's curvature, the range is, broadly speaking, limited to the visual horizon. It is for this reason that television transmitters are installed on tall masts or towers, which have to be spaced about fifty miles apart in order to provide good television coverage throughout a region (Fig. 6).

1. 1/30 second in U.S.A.

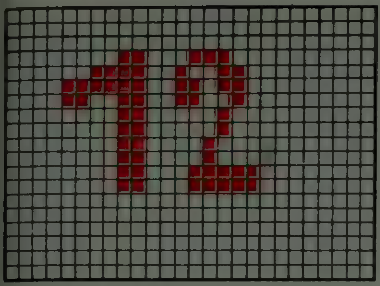


Fig. 1 PICTURE ELEMENTS

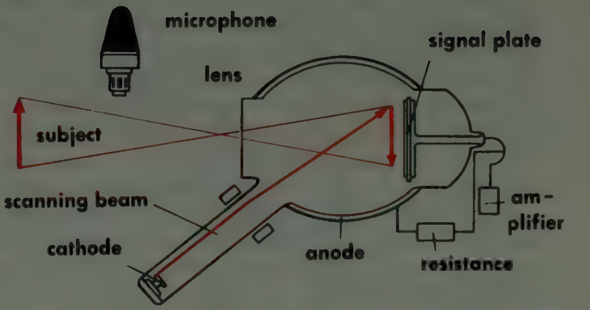


Fig. 2 TELEVISION CAMERA (iconoscope) AND MICROPHONE

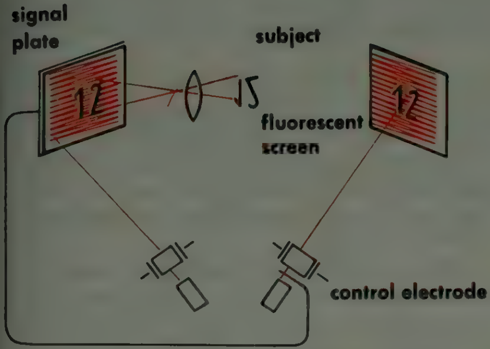


Fig. 3 ICONOSCOPE (transmission) AND CATHODE RAY TUBE (reception)

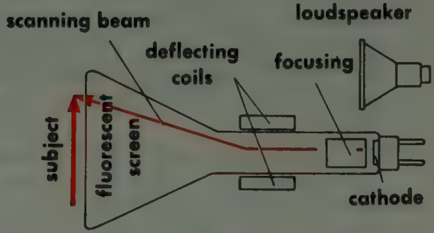


Fig. 4 SOUND AND PICTURE REPRODUCTION AT RECEIVER (loudspeaker and cathode ray tube)

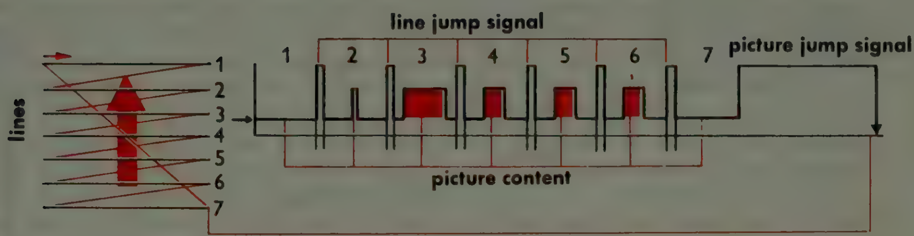


Fig. 5 SCANNING OF AN IMAGE AND CORRESPONDING ELECTRIC SIGNALS

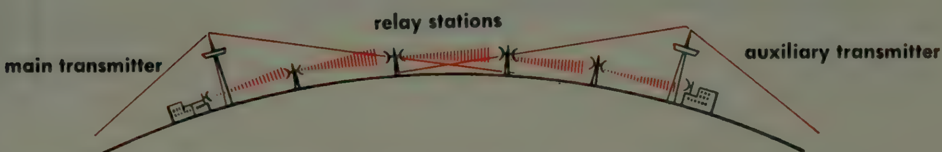


Fig. 6 TELEVISION TRANSMISSION WAVES TRAVEL IN STRAIGHT PATHS; RELAY STATIONS AND AUXILIARY TRANSMITTERS ARE THEREFORE REQUIRED

The *iconoscope*, invented some forty years ago and now obsolescent, is the oldest electronic analysing device for converting the optical image in the television camera into a sequence of electric signals. It proved to be much more efficient than the older mechanical scanning devices, which it entirely superseded in the nineteen-thirties. The essential component of the iconoscope is the signal plate, whose front face is covered with a mosaic consisting of hundreds of thousands of tiny globules of silver, which are so treated during manufacture that each globule has a surface of the oxides of silver and cesium. A lens focuses an optical image on the mosaic, the whole surface of which acquires a positive photo-electric charge whose distribution matches the distribution of light in the image (Fig. 1). The mosaic is scanned line by line by a narrow beam of electrons which is moved across the image. When this beam hits a globule, the latter is discharged. The individual silver oxide coatings act as one electrode of a condenser, the other electrode of which is the metallic signal plate (separated from the globules by a sheet of mica) (Fig. 2). Each time a globule is discharged, it gives an electrical impulse (Fig. 3). The translation of the televised scene into its electrical counterpart thus results in a sequence of electrical impulses known as the television picture signal.

The line-by-line, left-to-right, top-to-bottom dissection and reconstitution of television images is known as scanning. The agent which disassembles and reassembles the light values along each line is called the scanning spot (produced by a beam of electrons), and the path it follows is the scanning pattern. The spot is moved from left to right and then returned rapidly, while extinguished and inactive, from right to left. At the same time, the spot is moved comparatively slowly from top to bottom.

The first successor to the iconoscope was the *orthicon*, in which the mosaic is similar to that of the iconoscope, but is composed of squares of photosensitive material. The signal plate is formed by a transparent metal coating on the reverse side. A further development is embodied in the *image orthicon*, which is notable for its very high sensitivity to light. It is similar to the orthicon, but includes an additional electrical-imaging process and comprises an amplifier based on the phenomenon of electron multiplication. The electron multiplier increases the strength of the picture signal by several thousand times, whereby the very high sensitivity is obtained.

Another television camera tube is the *vidicon*, whose function is based on the phenomenon of photoconductivity. In its early forms its action was rather slow, and it was limited to industrial applications, traffic supervision, etc. Later developments overcame this drawback. The signal plate in the vidicon is a transparent metallic coating on which is a layer of photoconductive material (a complex compound of selenium) whose electrical resistance is high in the dark but diminishes as the light increases. The optical image induces a pattern of varying conductivity which corresponds to the distribution of brightness in the image. The conduction paths through the photoconductive layer allow positive charge from the signal plate to pass through the layer. An electron beam neutralises the positive charge on each point of the electrical image, and the resulting change in potential is transferred by "condenser" action to the signal plate (on the same principle as in the iconoscope). An advantage of the vidicon is that it can be constructed to a very small size, so that small and relatively inexpensive lenses can be used.

Fig. 1 ICONOSCOPE

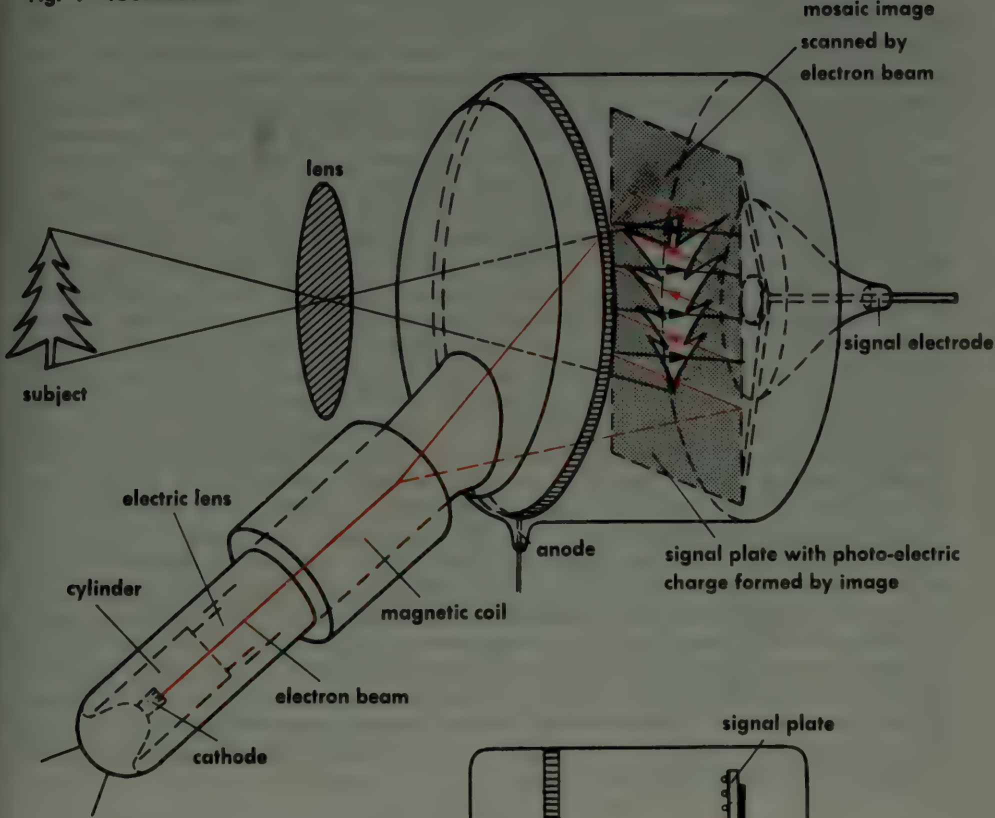
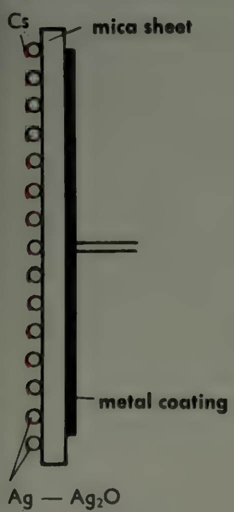


Fig. 2 SIGNAL PLATE



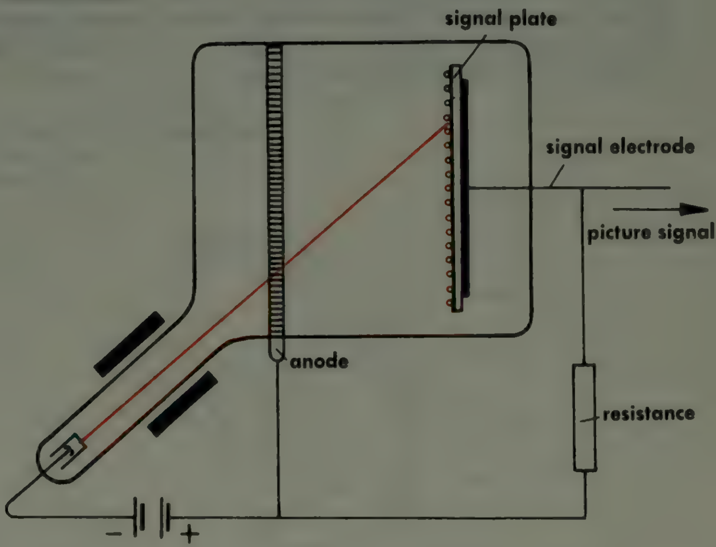


Fig. 3 ELECTRICAL CONNECTIONS OF ICONOSCOPE

The television picture tube (cathode-ray tube) is equipped with an electron gun comprising a hot cathode which emits electrons. These are concentrated into a beam which is moved to and fro by a deflecting system and appears as a spot of light on a fluorescent screen. Concentration of the emitted electrons into a beam is done by electrodes which function as an electric lens, or by a concentrating coil functioning as a magnetic lens (Fig. 2). Electrostatic deflection of the beam in two directions for scanning is obtained by means of two mutually perpendicular pairs of plates (electrodes) (Fig. 1) or by means of a magnetic deflecting coil (Fig. 2). The advantage of magnetic over electrostatic deflection is that larger deflection angles can be achieved at low voltages. This in turn enables the tube to be made shorter, so that the television set can be of "flatter" construction. The scanning spot of the electron beam moves to and fro across the screen, line by line, in synchronisation with the scanning spot in the television camera. The fluorescent screen, which is inside the tube, consists of a coating of chemicals (of which, for example, zinc sulphide may be a major constituent) which glow under the impact of high-speed electrons. The colour of the fluorescence can be modified by certain admixtures to the coating. The concentration of the electron beam in the "electric lens" is achieved by means of the electric field which is formed between the earthed cylinder (zero potential) and the anode plate (positive potential). The lines of force pass through the hole at the centre of the anode, and the equipotential surfaces are curved rather in the manner of an optical lens (cf. page 150 and page 166). With magnetic concentration the electrons travel along spiral paths (cf. page 92). Scanning by electrostatic deflection of the electrons in the field of the two pairs of deflecting plates is comparable to the fall of a body in a gravitational field. Each individual electron describes a parabolic path. In the case of magnetic deflection a similar effect is achieved by causing the electrons to travel along a spiral path for some distance. Magnetic scanning is accomplished by two sets of electromagnet coils wound on a core of magnetic material. Deflection of the beam occurs by virtue of the fact that an electron in motion through a magnetic field experiences a force at right angles both to its direction of motion and to the direction of the magnetic lines of force.

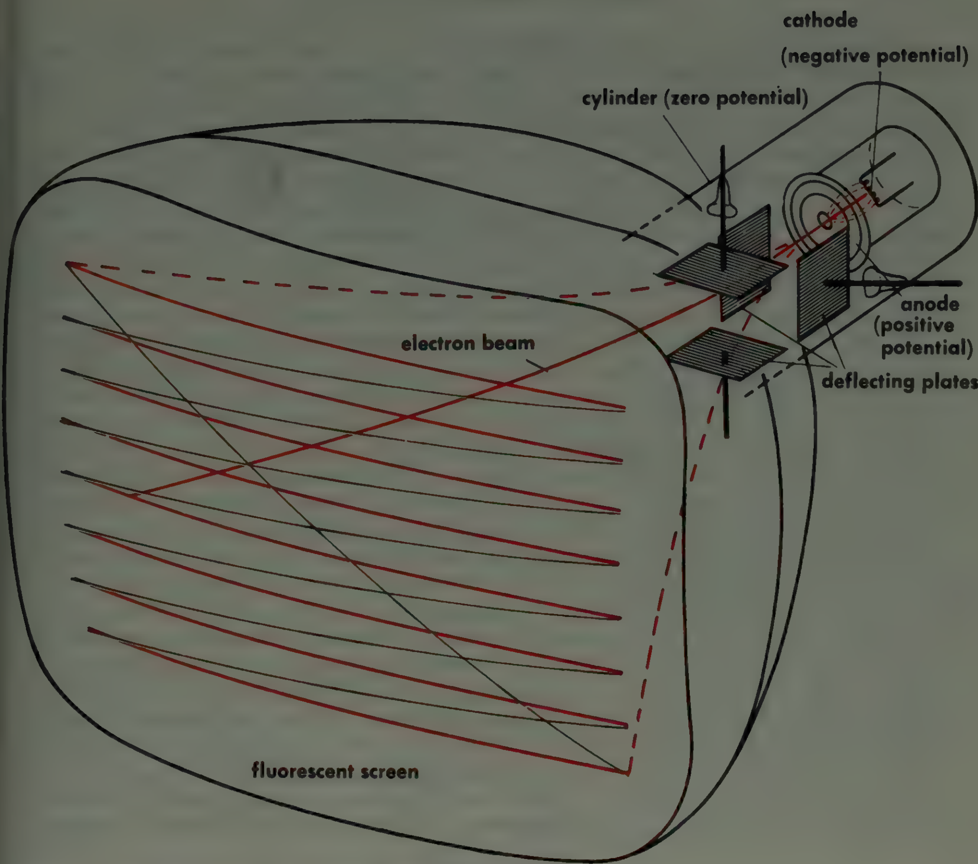


Fig. 1 ELECTROSTATIC FOCUSING AND DEFLECTION IN THE CATHODE RAY TUBE
(television picture tube)

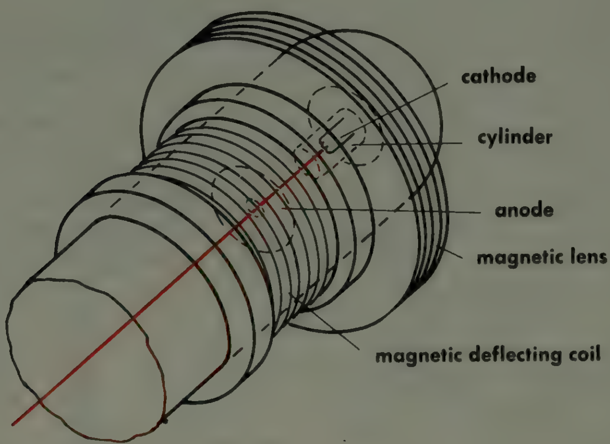


Fig. 2 MAGNETIC FOCUSING AND DEFLECTION DEVICES

An image converter tube is an apparatus which converts images of an optically non-observable kind of radiation into images which come within the range of optical observation. In principle (Fig. 1) it consists of a thin photocathode which reacts to invisible radiation. The electron image which is formed is projected on to the fluorescent screen by means of magnetic or electric focusing lenses (Figs. 2 and 3) and thus produces an optically visible image. For the conversion of X-rays a converter tube is used which contains a thin aluminium foil (Fig. 4) one side of which is provided with the fluorescent screen, while the other side constitutes the photocathode. With the aid of the X-ray image converter tube the original silhouette images can be electronically intensified. In this way, for instance, a sufficiently bright image can be obtained even with a low radiation dose which is quite harmless to the patient.

The mode of functioning of the image converter tube is based on the properties of electric and magnetic lenses (cf. page 166) whereby electron beams emitted from an (electronic) image point can be gathered and focused at another point. The concept of "lens" is derived from conventional optics and refers to devices whose only features in common with optical lenses are that they are able to focus rays, though in this case these are electron rays, not light rays. With the electric lens the focusing action is obtained by means of an electrical field (e.g., between two coaxial cylindrical electrodes, one of which has a higher potential than the other) which has curved equipotential surfaces (surfaces of constant potential or voltage) that are comparable to the curved surfaces of a lens (Fig. 2). Each individual electron performs vibrations of diminishing amplitude as it travels along in the direction of the cylinder axes. The electron beam, consisting of vast numbers of electrodes, thus assumes a tapering tubular shape, i.e., it converges and is focused. The functioning of a magnetic lens is even less like that of an ordinary optical lens (Fig. 3). In the longitudinal magnetic field, which is approximately parallel to their direction of flight, the electrons describe spiral paths. All electrons which start from one and the same point may travel along different spiral paths, but will eventually converge at one point. In this sense there is a focusing—i.e., image-forming—effect. Magnetic lenses are preferable to electrical ones because they can be operated at lower and therefore less dangerous voltages.

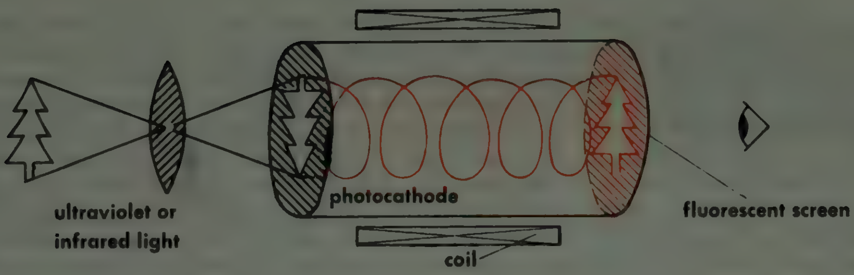


Fig. 1 PRINCIPLE OF IMAGE CONVERTER TUBE

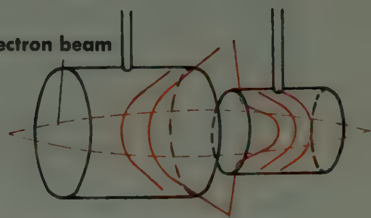


Fig. 2 ELECTRIC LENS

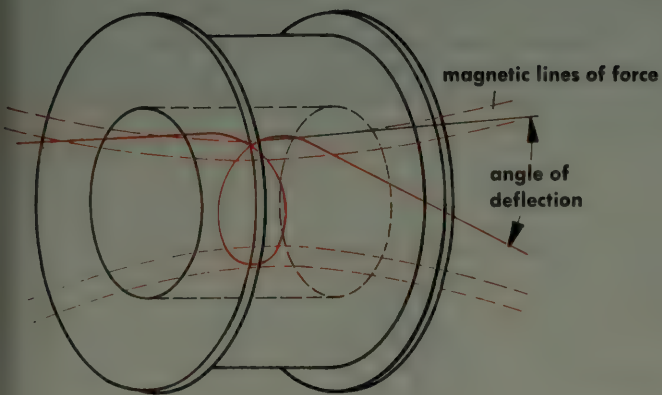


Fig. 3 MAGNETIC LENS

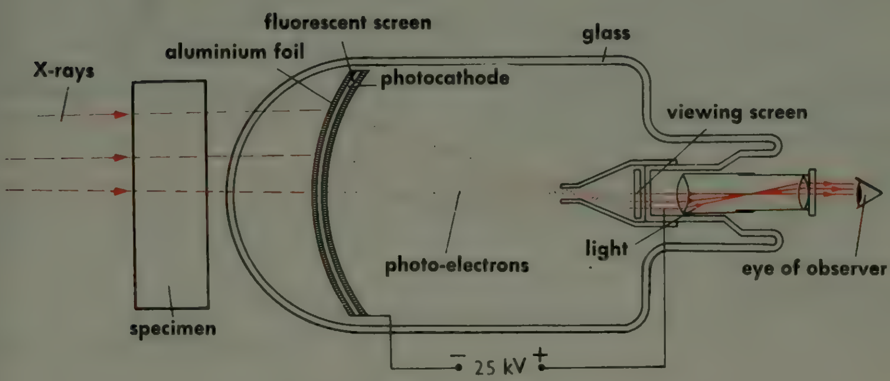
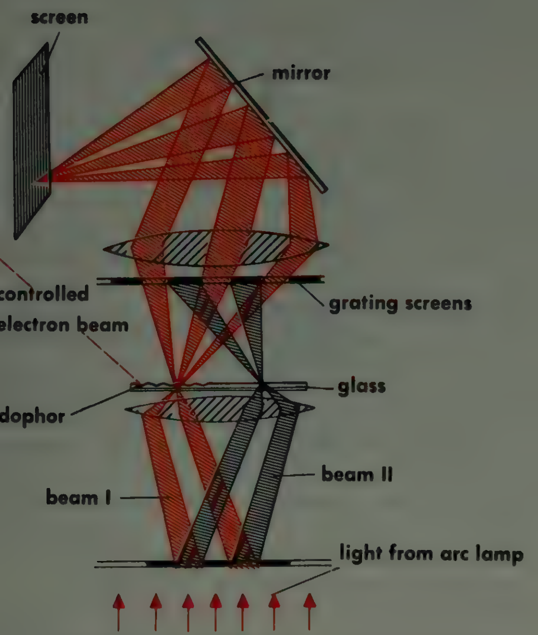


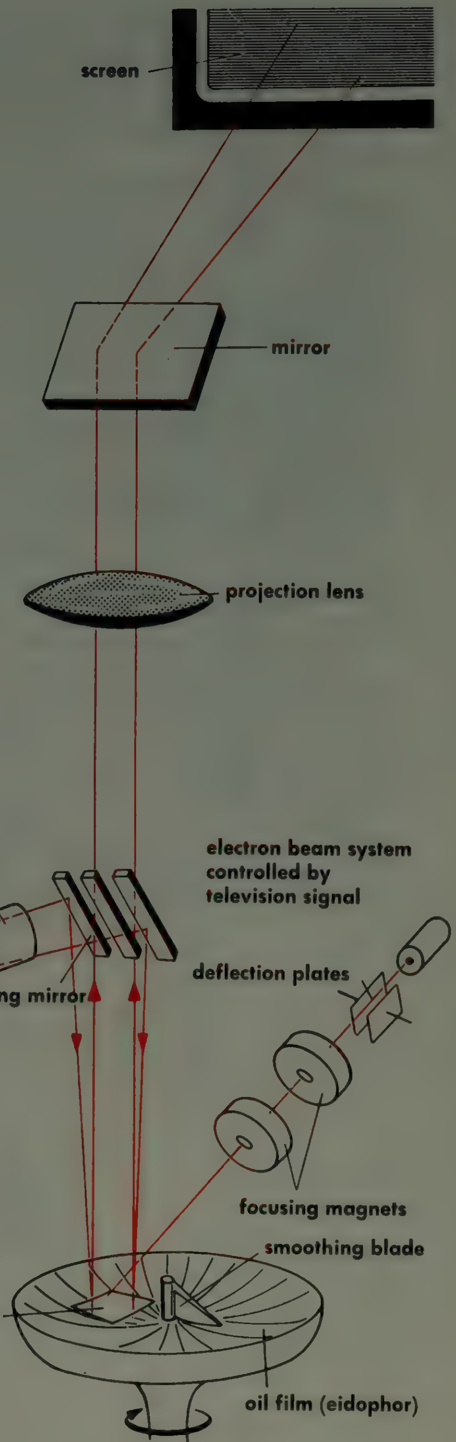
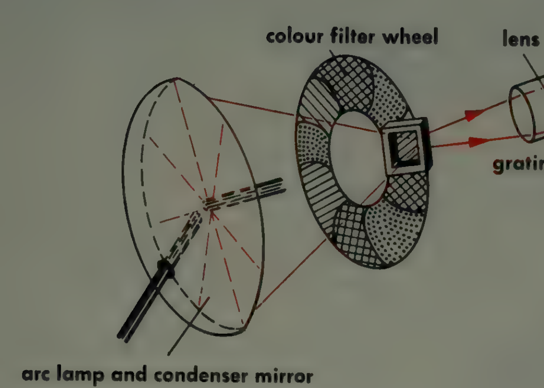
Fig. 4 X-RAY CONVERTER TUBE (schematic)

The eidophor system is a projection television system, i.e., it enables enlarged television pictures to be projected on to a screen. An ordinary television picture (see page 144) derives its brightness from the fluorescence of a screen bombarded by electrons; with the eidophor system, on the other hand, a very powerful source of light is controlled by the television signal picked up by the receiver. The light emitted by an arc lamp (which in the case of colour television is passed through a colour filter disc) is directed on to the screen by means of two grating screens (Fig. 1)—or, alternatively, a grating mirror silvered on the back (Fig. 2)—through an oil film and a projection lens. The oil film is given an electric charge of varying intensity, depending on the brightness value of the incoming signal, by an electron beam which is controlled by this signal. As a result of the electrostatic repulsion forces, the oil film thereby acquires varying degrees of curvature. The curved surface produces a change in the optical image formed by the light which passes through or is reflected by it. The light reflected by the bars of the grating mirror will, on its return path after being reflected by the oil film, pass with varying intensity through the gaps in the mirror, so that the stream of light is controlled by the television signal. In consequence of the scanning by means of the electron beam an electrical charge pattern is formed which corresponds to the brightness values of the image. This varying pattern of electrical charge in turn controls the light emitted by the light source which projects the greatly enlarged image on to the screen. The light rays which are unaffected by the charge pattern are retained by the grating screen (Fig. 1) or by the grating mirror (Fig. 2).

The apparatus represented diagrammatically in Fig. 2 must, of course, be accommodated in a high-vacuum enclosure to enable electronic control to be effected. The electrical charge of the oil film dies away gradually; for this reason the film is made to rotate slowly away from under the controlling electron beam, and a smoothing blade ensures that the film is given a fresh, electrically discharged surface before it is again exposed to the electron beam.



. 1 EIDOPHOR PROJECTION SYSTEM (schematic)

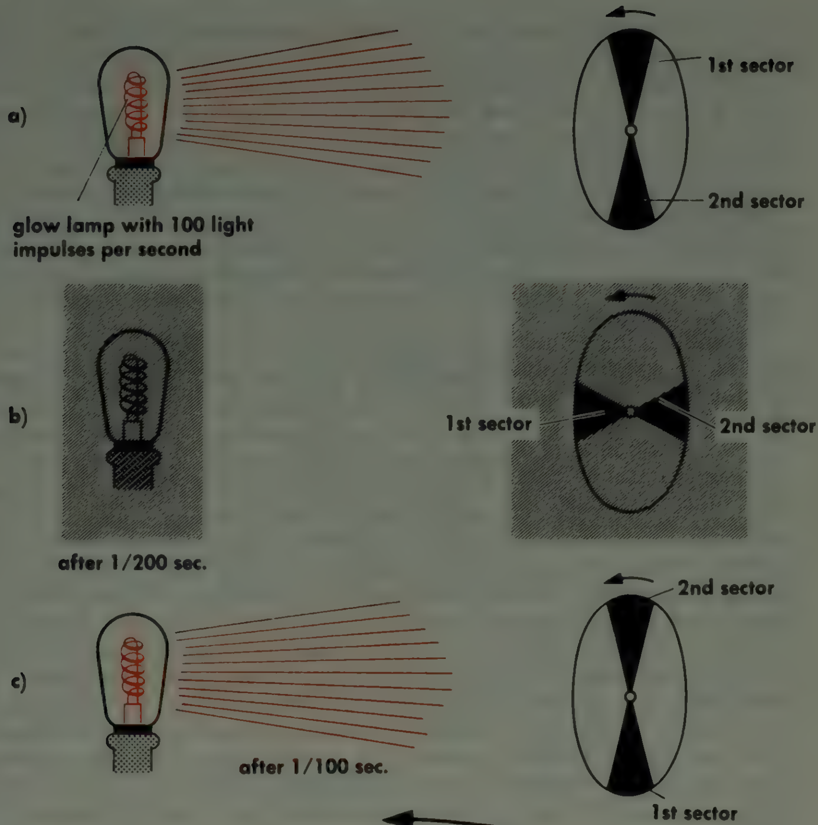
ig. 2 EIDOPHOR EQUIPMENT

Stroboscopic effects have a physiological basis. The human sense of vision is so slow to react to light stimuli that it is unable to separate two different light impulses reaching the eye within a very short period of time (less than $\frac{1}{10}$ second). A succession of impulses following one another at such brief intervals are observed by the eye as a continuous unbroken sequence (principle of cinematography and television; cf. Figs. 2 and 3).

Stroboscopic phenomena in a more restricted sense occur in cases where two periodic motions which differ slightly in their timing, or which are performed synchronously, are superimposed one upon the other. An example is provided by the glow lamp in Fig. 1, which is fed with alternating current (50 cycles/sec.) and emits 100 light impulses (brief flashes) every second. This intermittent light is beamed on to a revolving disc provided with two dark-coloured sectors. If the disc is rotated exactly 180° (half a revolution) during the interval of darkness (Fig. 1b) between light maxima (Figs. 1a and 1c), the observer will see a stationary two-pointed star. The speed of rotation of the disc is then 50 revolutions per second. If the speed is a little below this value the double sector will appear to be rotating in a direction opposite to the direction of the rotation of the disc; if the speed of the disc is a little above 50 revolutions per second, the double sector will appear to be rotating in the direction of rotation of the disc. In the former case the disc has not quite rotated through 180° when the next light maximum occurs; in the latter case the disc has rotated a little more than 180°. Such stroboscopic effects are sometimes seen in the cinema when the frequency with which the successive images are projected on to the screen is superimposed on the speed of rotation of, for example, the spokes of a wheel. When that happens, the wheel appears to be revolving backwards.

Examples of stroboscopic phenomena in the more general sense are afforded by the toy called a stroboscope (Fig. 2) or the little books with pictures which appear to move when the pages are flicked (Fig. 3). In the stroboscopic toy a succession of pictures, each corresponding to a stage of movement, is viewed through rotating slots, with the result that an impression of continuous motion is obtained. In engineering, stroboscopes are used for measuring and checking the speeds of rotation of shafts and other parts of machinery; they are also used for apparently slowing down periodically repetitive motions and thus enable them to be observed more conveniently.

Fig. 1



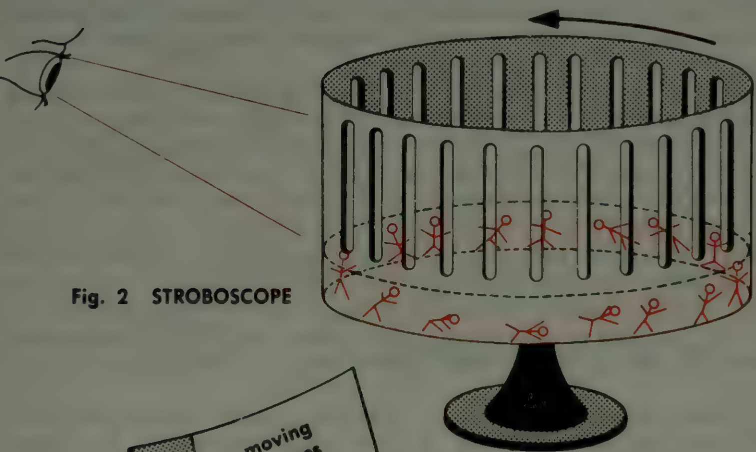


Fig. 2 STROBOSCOPE

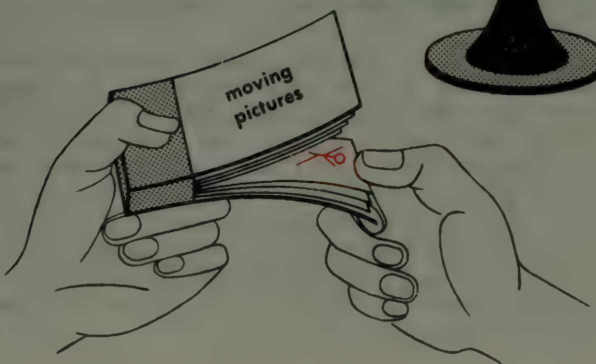


Fig. 3

When a ray of light strikes the boundary surface of two different transparent media, e.g., on passing from air into a block of glass, a proportion of the light is reflected back into the first medium. The rest penetrates into the second medium, but it undergoes a change of direction in doing so: it is refracted. These two phenomena of reflection and refraction are indicated in Fig. 1. The ray r is so reflected that the angle α which it forms with the perpendicular of incidence l (i.e., a line perpendicular to the boundary surface) is equal to the angle α formed by the incident ray e (angle of incidence = angle of reflection). The incident ray e, the perpendicular of incidence l, the reflected ray r, and the refracted ray g are all situated in the same plane. The mathematical relation that exists between the angle of incidence α and the refraction angle β is known as the law of refraction: $\sin \alpha / \sin \beta = n$. This ratio n is called the index of refraction.

On undergoing refraction, white light is broken up into a number of different colours, the spectral colours (Fig. 2): a spectrum is formed (Fig. 3). Red light undergoes the least, violet light undergoes the greatest amount of diffraction (dispersion). If the light rays corresponding to the various spectral colours are combined by means of a lens, white light is obtained again. The same result can even be obtained with only two spectral colours (in a certain intensity ratio) if these are complementary to each other (complementary colours). Each spectral colour corresponds to a particular wavelength of light. The spectrum of an incandescent solid or liquid substance or a very highly compressed gas is continuous: it contains all wavelengths (comprising "all the colours of the rainbow" from red to violet). The fact that particular wavelengths are nevertheless absent from the solar spectrum (there are individual dark lines, known as Fraunhofer lines) is due to absorption of these wavelengths in the outer, gaseous part of the sun. Under normal conditions, incandescent gases emit light of particular, definite wavelengths; the spectrum takes the form of a line spectrum, consisting of individual spectral lines (in the case of atoms), or a band spectrum (in the case of molecules) which comprises a large number of spectral lines spaced close together. These wavelengths can similarly be absorbed by the gas. Every chemical element, when it is in the gaseous state, produces a characteristic spectrum whereby it can be detected (spectrum analysis; Kirchhoff and Bunsen, 1859). A spectroscope is used for the purpose (Fig. 4).

If an object of sufficiently small dimensions is placed in the path of light rays, it does not cast a sharply outlined geometrical shadow (as would correspond to rectilinear propogation of light). Instead, we observe a certain arrangement of light and dark areas (e.g., stripes). Such deviation from the rectilinear path is called diffraction. This phenomenon can be explained only in terms of the wave theory of light: superposition of individual light waves (interference). An arrangement of very narrow, closely spaced gaps or slits forms a so-called diffraction grating (Fig. 5). If white light is directed on to a diffraction grating, it will likewise produce a spectrum, because the directions in which the bright areas occur are dependent on the wavelength. However, in contrast with what happens in refraction, in diffraction the red light undergoes a greater deflection from its original direction than does the violet light with its shorter wavelength.

The small annular haloes, or coronas, which sometimes appear round the sun and moon are diffraction phenomena caused by tiny water droplets in the upper atmosphere. For this reason the outer edge of such a halo displays a (faint) red colour. On the other hand, large haloes are phenomena due to refraction and reflection of light by ice crystals (usually in the form of cirro-stratus clouds). The colour sequence in such haloes, if it is perceptible at all, is the reverse of that in small haloes, i.e., red is on the inside and violet on the outside of the ring. A rainbow is the result of a combination of refraction, reflection and diffraction phenomena. The blue colour of the sky is due to scattering of sunlight on its way through the atmosphere, in which process the shorter wavelengths (blue light) are more strongly scattered than the longer wavelengths (red light).

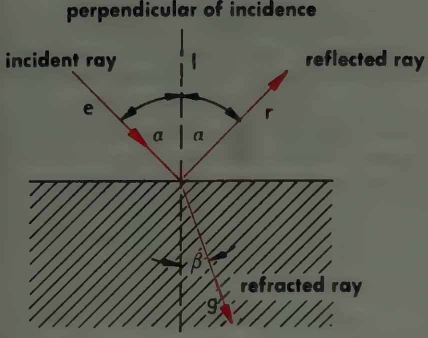


Fig. 1 REFLECTION AND REFRACTION

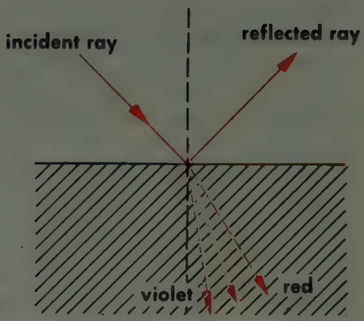


Fig. 2 BREAKING-UP OF LIGHT ON REFRACTION

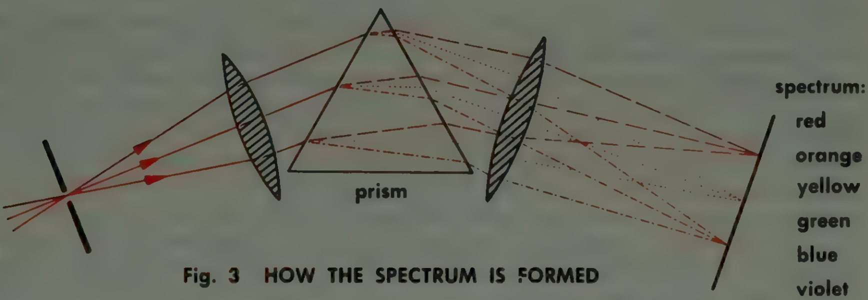


Fig. 3 HOW THE SPECTRUM IS FORMED

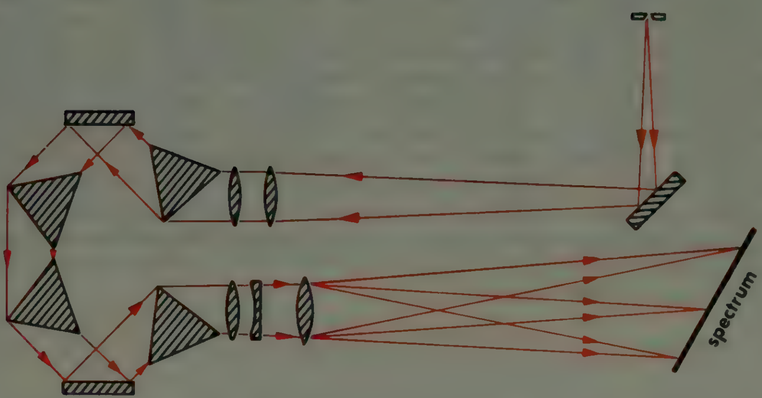


Fig. 4 PRINCIPLE OF A SPECTROSCOPE

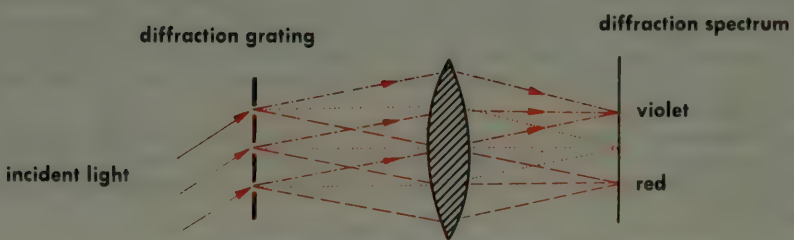


Fig. 5 HOW A DIFFRACTION SPECTRUM IS FORMED

When rays of light pass through a prism, they undergo a change of direction: they are always deflected away from the refractive edge (see page 156). It is possible to conceive an assembly of prisms, arranged as shown in Fig. 1a, whose respective refractive surfaces progressively become more nearly parallel to each other towards the middle: light rays passing through the outer prisms will undergo the greatest amount of refraction, with consequent deflection of their path towards the centre, whereas the middle prism with its two parallel surfaces causes no deflection at all. When a beam of parallel rays passes through these prisms, the rays are all deflected towards the axis and converge at one point (F'). Rays emerging from a point P are also so deflected by the prisms that they converge at a point P'. A lens can be conceived as consisting of a large number of such prisms placed close up against one another, so that their surfaces merge into a continuous spherical surface. A lens of this kind, which collects the rays and concentrates them at one point, is called a convergent lens. Since it is thicker in the middle than at the edge, it is known as a *convex lens.*

In the case of a *concave lens*, which is thinner in the middle than at the edge, similar considerations show that all rays diverge from the centre (Fig. 1b). Hence such a lens is called a divergent lens. After undergoing refraction, parallel rays appear to come from one point (F), while rays emerging from a point will, after passing through the lens, appear to emerge from another point (see Fig. 3b). Fig. 2 shows the various shapes of convex and concave lenses. The last lens in each group has surfaces curved in the same direction but having different radii of curvature; these are known as meniscus lenses and are used more particularly in spectacles (see page 162).

The properties of lenses are determined by the law of refraction: Rays parallel to the axis of a convex lens are refracted by the lens so as to converge at a point F, called the focus or focal point. Conversely, rays which pass through (or emerge from) become parallel after being refracted by the lens. Rays which pass through the centre of the lens continue with their direction unchanged. In the case of a concave lens, rays parallel to the axis are refracted in such a manner that they appear to emerge from a single point, which here again is called the focus. With the aid of these optical properties of lenses it is possible to construct the path of the rays and determine the position of the image that the lens will form of any particular point of an object (Figs. 3a and 3b). A convex lens generally produces a *real image* which is formed at the point of convergence of the rays and can be made visible by projecting it on to a screen. On the other hand, a concave lens produces a *virtual image*, from which the light rays appear to diverge. Such an image cannot be formed on a screen.

The positions of the image and the object in the case of a thin lens are determined by the following simple formula:

$$\frac{1}{g} + \frac{1}{b} = \frac{1}{f} = D$$

(g = object distance; b = image distance; f = focal distance or focal length). Instead of the focal length, its (refractive) power is expressed in dioptres. The shorter the focal length of a lens is, the greater is its power. A convergent lens of 1 metre focal length is said to have a power of +1 dioptre. The power of a convergent lens of 0.2 m focal length is +5 dioptres. The power of a divergent lens is expressed as a negative quantity; the focal distance of such a lens is likewise negative, which means that b must also have a negative value, as the virtual image is located on the same side of the lens as the object is.

Single lenses are affected by errors due to physical causes which prevent truly accurate convergence of the light rays. The sharpness of the image can be improved by using a combination of two or more lenses whose errors compensate one another (see page 178).

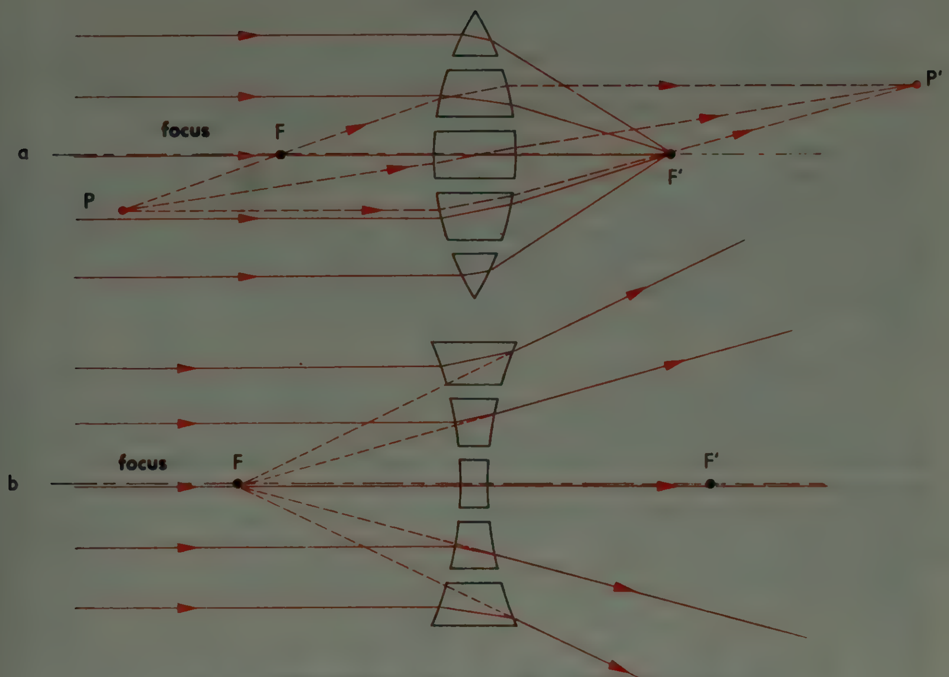


Fig. 1 CONVERGENT LENS (*a*) AND DIVERGENT LENS (*b*)

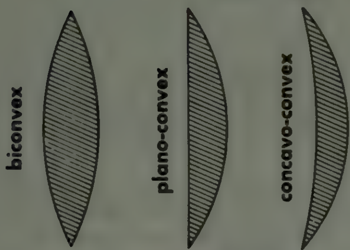


Fig. 2a CONVERGENT LENSES

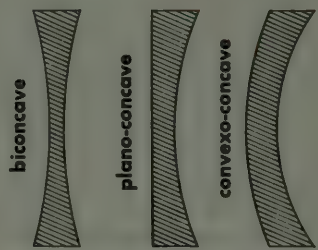


Fig. 2b DIVERGENT LENSES

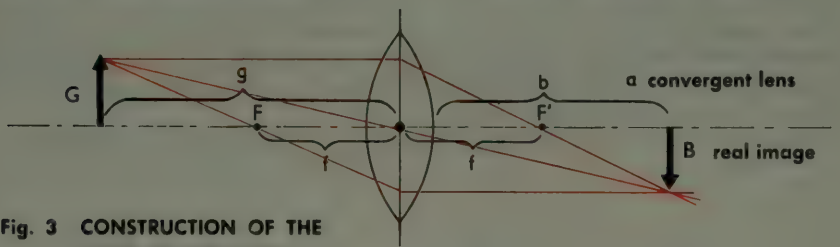
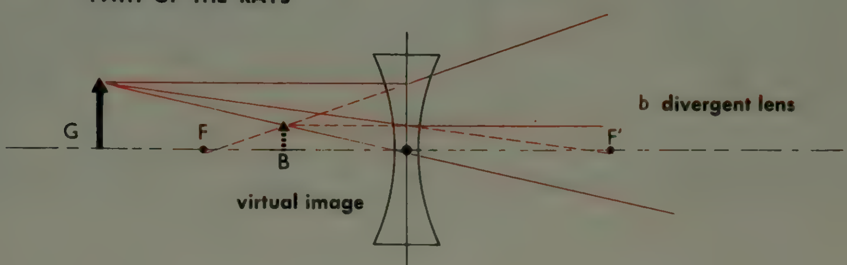


Fig. 3 CONSTRUCTION OF THE PATH OF THE RAYS



Reflection occurs at smooth surfaces. The law of reflection states: the incident ray, the perpendicular to the reflecting surface, and the reflected ray are situated in one plane. The angle of reflection is equal to the angle of incidence (Fig. 1).

Plane mirror

Rays emerge from the object in all directions. If a beam of rays from the object is reflected by a plane mirror, the rays will change their direction but will, after doing so, continue on their divergent paths. They do not converge and therefore cannot produce a real image. They all appear to come from a point located behind the mirror, i.e., from the virtual image of the object. The virtual image is at the same distance behind the reflecting surface as the object is in front of it.

Curved mirror

If a spherical reflecting surface is conceived as being composed of a vast number of tiny plane surfaces, of which each is perpendicular to a radius of the sphere, then the law of reflection can be applied to such a mirror. Obviously, a ray of light directed through the centre of curvature M will be reflected back in its own direction. At each point the reflected ray forms the same angle with the radius as does the incident ray, so that rays parallel to the principal axis will converge at one point located midway between the mirror and its centre of curvature. This point of convergence is the focus or focal point (F) of the concave mirror. The focal length of such a mirror is equal to half the radius of curvature. Conversely, all rays passing through, or emerging from, the focus will be reflected in a direction parallel to the principal axis. Fig. 2 shows the construction of the ray paths associated with a convergent (concave) mirror. The formula linking the positions of the object and image is similar to that already given for lenses:

$$\frac{1}{g} + \frac{1}{b} = \frac{1}{f} = \frac{2}{r}$$

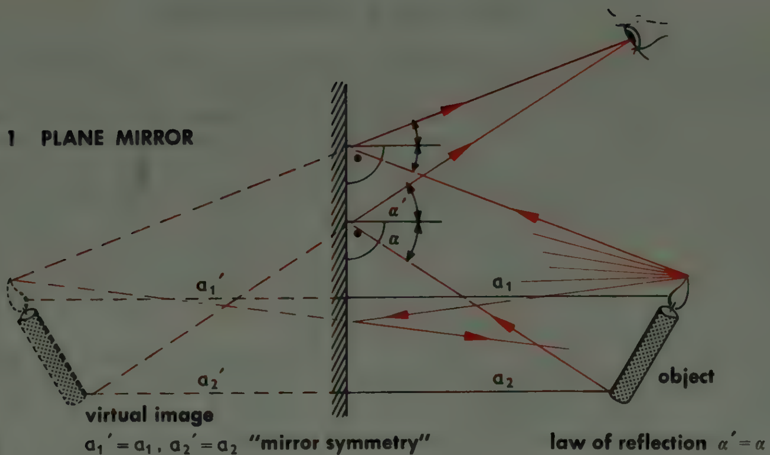
(g = object distance; b = image distance; f = focal distance or focal length = $\frac{1}{2}r$; and r = radius of curvature). In the case of a divergent (convex) mirror a virtual image is formed behind the mirror; the focal length and image distance then both occur as negative quantities in the above equation.

Driving mirrors in motor cars are sometimes of the convex type (example 1), forming a virtual image of reduced size but covering a wide field of vision behind the vehicle. A shaving mirror (example 2) is a concave mirror. The user comes within its focal distance, with the result that he sees the upright and enlarged virtual image of his face in the mirror.

The constructions for determining the position of the image carried out in the accompanying drawings are in each case based on the choice of two particularly convenient rays, namely, the ray parallel to the principal axis and the ray passing through the focal point of the mirror.

The use of mirrors in the reflecting telescope is explained on page 170.

Fig. 1 PLANE MIRROR



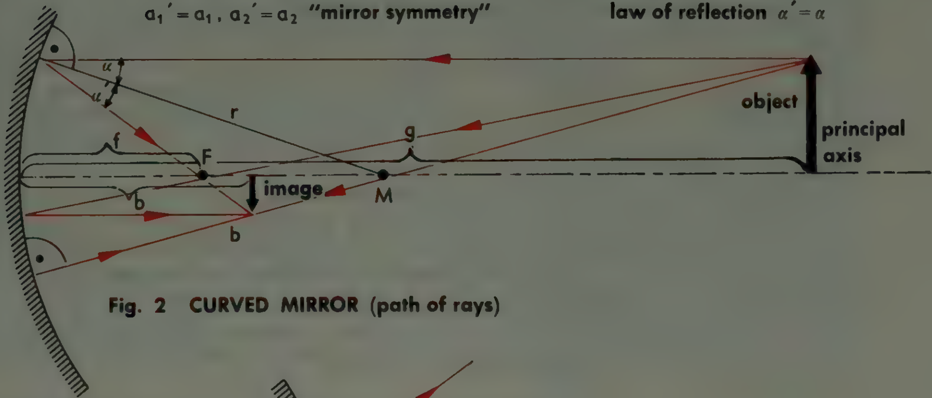
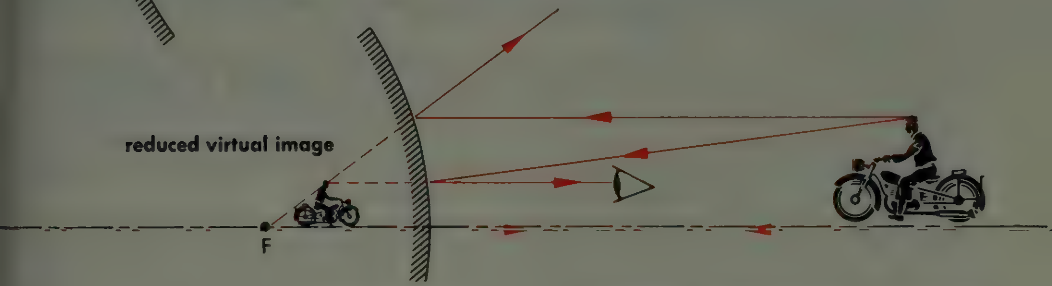
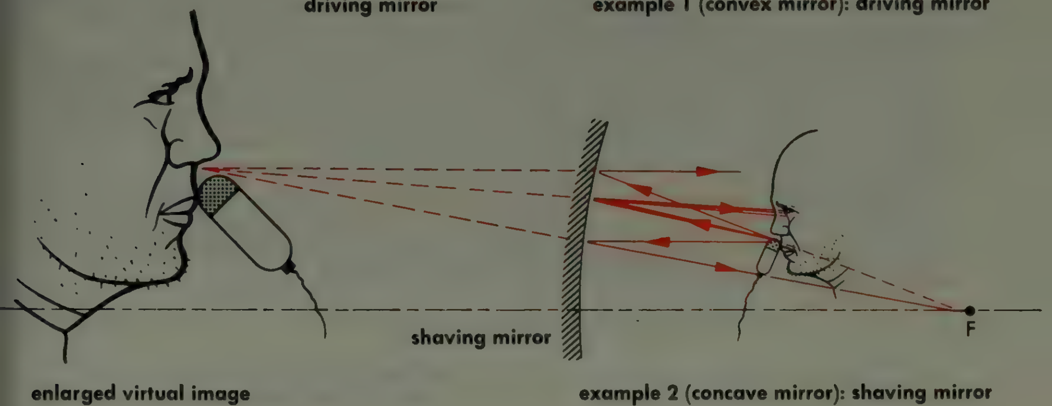


Fig. 2 CURVED MIRROR (path of rays)



example 1 (convex mirror): driving mirror



example 2 (concave mirror): shaving mirror

When a person has normal eyesight, the image of a distant object (theoretically at infinity) is formed accurately on the retina of the eye, i.e., the focus (see page 158) is located on the retina (Fig. 1). Since the image distance—the distance from lens to retina—is predetermined by the size of the eyeball, the image of an object nearer the eye can be sharply focussed on the retina only by reduction of the focal length of the lens, by increasing its refractive power (see page 158). This is achieved by the action of a muscle, which increases the curvature of the lens. When a distant object is observed, the radius of curvature of the front surface of the lens is about 10 mm (0.4 in.). To adjust the focus to an object only 4 inches away, this radius of curvature must be reduced to about half this amount. This adjustment of the eye to varying distances is called "accommodation" (Fig. 2).

However, not all human eyes are "normal" in that they behave as represented in Figs. 1 and 2. If the distance between the lens and the retina is too large, the image of a distant object will be formed not on, but in front of, the retina (Fig. 3). To form a sharply focused image, the object must be at a shorter distance from the lens, i.e., the eye is only able to form a clear image of objects within this latter distance. Such a person is said to be *short-sighted*. This defect can be corrected by spectacles (Fig. 3). Light rays coming from infinity are made slightly divergent by the lens, so that it is as if they come from an object situated closer to the eye. By thus interposing a lens with negative refractive power between the eye and the object, the combined system of glass lens and eye lens has its focus located on the retina.

If the eyeball is too short for the focal length of the lens, the person is said to be *far-sighted* (Fig. 4); his eyes must accommodate to get even a distant object properly in focus (whereas a normal eye lens is "at rest", i.e., does not have to accommodate, when viewing objects at infinity), while near-by objects cannot be seen sharply at all. This condition is corrected by means of spectacles fitted with convergent lenses. The combined system of glass lens and eye lens is again so contrived that the focus is on the retina, so that distant objects can be viewed without the strain of constant accommodation.

The amplitude of accommodation of the human eye, i.e., the limits of distance within which an object must lie in order to enable a sharp image to be formed on the retina, diminishes with age. This is because the lens grows less flexible, so that its curvature becomes more difficult to vary. A person with normal eyesight may, at the age of 55, be unable to bring objects less than about 3 ft. away sharply into focus, and he will require reading-glasses (compare Fig. 5 with Fig. 2).

Loss of power of accommodation can to some extent be compensated by spectacles fitted with bifocal lenses. These act as weak lenses when the gaze is directed straight ahead and as strong ones when the gaze is directed downward for reading or working. Trifocals are also sometimes used. The accommodation amplitude of the eye and the refractive power of spectacle lenses is measured in dioptres (see page 158).

In the defect known as astigmatism the refracting surfaces of the eye have unequal curvature, which prevents the focusing of light rays to a common point on the retina. Correction is achieved by means of spectacles whose lenses embody a combination of spherical and cylindrical curvature, so contrived that the combined system comprising glass lens and eye lens has the correct spherical curvature.

Contact lenses, as an alternative to conventional spectacles, are lenses worn on the eyeball. All visual conditions correctible by spectacles can also be corrected by contact lenses. These were first introduced towards the end of the previous century and were made of glass. Today they are made of plastics.

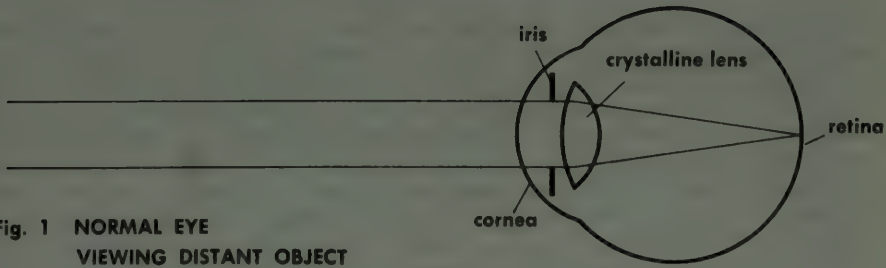


Fig. 1 NORMAL EYE
VIEWING DISTANT OBJECT

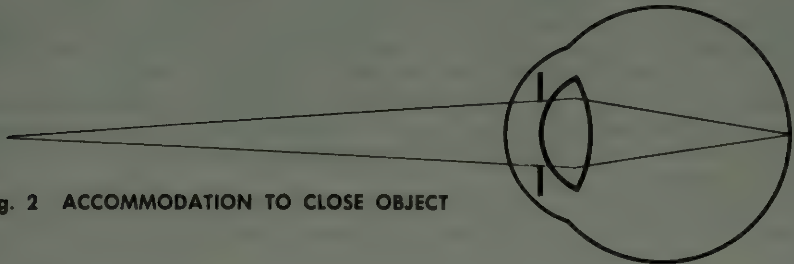


Fig. 2 ACCOMMODATION TO CLOSE OBJECT

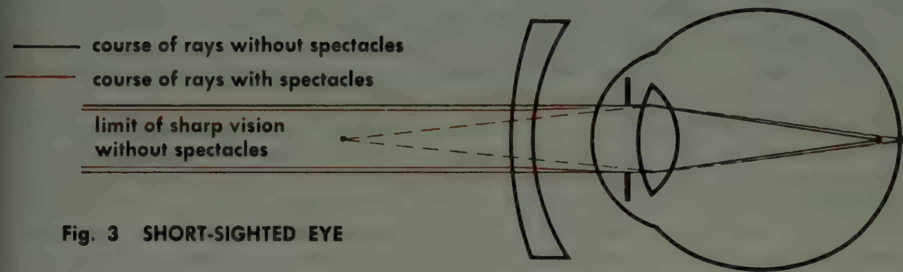


Fig. 3 SHORT-SIGHTED EYE

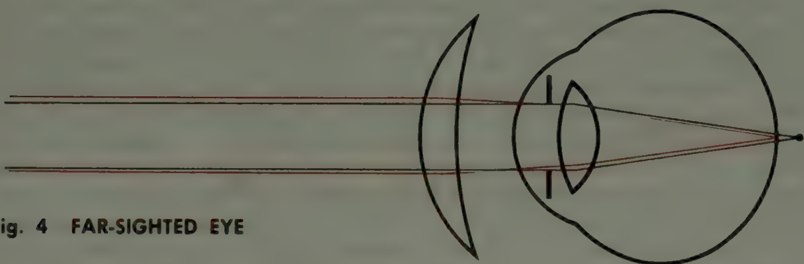


Fig. 4 FAR-SIGHTED EYE

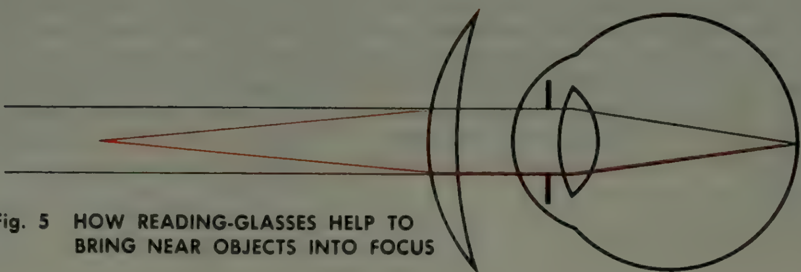
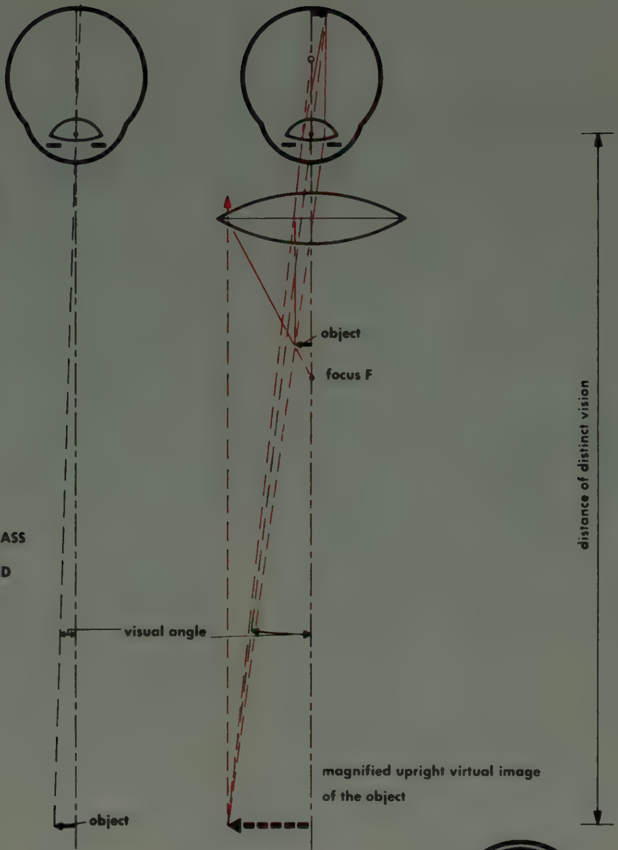


Fig. 5 HOW READING-GLASSES HELP TO
BRING NEAR OBJECTS INTO FOCUS

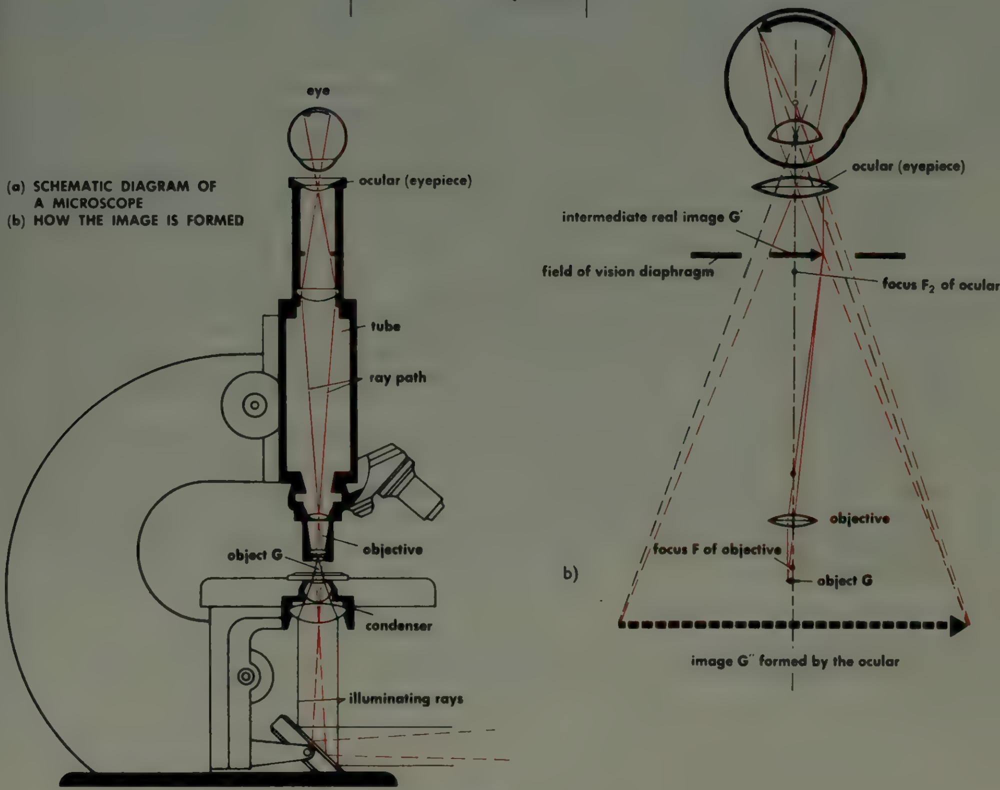
How clearly the human eye can distinguish details on any particular object depends on the visual angle under which they are observed. Distant objects appear small, and to see them better we use a telescope, which increases the visual angle (see page 170). In the case of very small objects the visual angle remains too small even when they are viewed from the shortest possible distance at which they can be brought into sharp focus by the unaided eye (approx. 6–15 in.). If an object is brought still closer to the eye, the visual angle does indeed increase, but the eye is then no longer able to see it distinctly. With the aid of a magnifying glass or a microscope, however, it is possible to bring the object very close to the eye and yet view it as though it were comfortably within the eye's range of accommodation (i.e., the range of distance within which it can produce a sharply focused image on the retina). The "magnifying power" of a microscope or magnifying glass is the ratio of the apparent size of the image of an object formed by the instrument to that of the object seen by the naked eye. For the purpose of this definition it is assumed that the object would be examined by the naked eye at the least distance of distinct vision, conventionally assumed to be 25 cm (about 10 in.). Thus, for instance, a microscope with a magnifying power of 300 will show an object 300 times as large as it would appear to the naked eye from a distance of 25 cm.

The *magnifying glass* (Fig. 1) consists of a convergent lens (or a convergent system of lenses) of short focal length. The object to be examined is placed within the focal distance, so that a magnified upright virtual image is formed approximately at a distance at which the eye can see it most distinctly. A magnifying glass is similar to reading glasses (see page 162) in that it enables the eye to take a close look at objects while remaining accommodated to a greater distance.

In a *microscope* (Fig. 2) magnification takes place in two stages. In the case of a telescope the object is a long distance away; the real image formed by the object lens (which has a long focal length) is formed approximately at the rear focal point. On the other hand, in the case of a microscope the objective (object glass) has a very short focal length (ranging approximately from $1\frac{3}{4}$ in. to 0.06 in.); the object is placed so close to the front focal point of this lens that the image distance is much greater than the focal length. As a rule, the latter is predetermined by the length of the microscope tube, which is usually about $6\frac{1}{2}$ in. Hence the real intermediate image (for the above-mentioned focal lengths) is about $2\frac{1}{2}$ to 100 times as large as the object viewed. In the plane of the intermediate image are the field of vision diaphragm and, in some microscopes for special purposes, measuring scales, measuring grids, or similar optical aids. The intermediate image is examined through a magnifying glass, the ocular (eyepiece), which is mounted at the top of the tube and which further magnifies the image. For example, the objective of a microscope has a magnifying power of 40 × and the ocular has a magnifying power of 10 × ; in that case the overall magnification will be 400 ×. With an objective and ocular of 100 × and 25 × magnifying respectively, a magnification of 2500 × would be obtained. However, an ordinary optical microscope of such power is of little practical value. Because of the wave character of light, it is possible only to distinguish details of a size down to approximately the wavelength of light (0.0004 – 0.0007 mm = 0.000016 – 0.000028 in.). For practical purposes this limits the maximum useful magnification to about 2000 ×. Much higher magnifications are attainable with the electron microscope, which uses electron rays of much shorter wavelength than visible light rays (see page 166). An important part in connection with the satisfactory functioning of the microscope is played by the condenser, i.e., the lens or lenses which serves to direct the light on to the object and to illuminate it intensively and uniformly in such a manner that nearly all of this light is transmitted into the object glass of the microscope.



. 1 THE PURPOSE OF A MAGNIFYING GLASS IS TO INCREASE THE VISUAL ANGLE UNDER WHICH THE OBJECT IS VIEWED



. 2 (a) SCHEMATIC DIAGRAM OF A MICROSCOPE
(b) HOW THE IMAGE IS FORMED

As the resolving power of the light microscope is limited by the wavelength of light (see page 164), efforts were made to utilise rays of shorter wavelength which can also be deflected and be used to form images. This possibility is presented by electron beams, i.e., free electrons which are accelerated to high velocities on traversing an electric field. Depending on the velocity of the electrons, such beams can be considered to have a certain wavelength; under particular conditions they behave as though they were of an undulatory character, i.e., composed of waves. As an electron has a negative electric charge, it undergoes acceleration on passing through an electric field (Fig. 1)—for example, it is attracted towards the positive plate of a condenser or, in general, it is accelerated in the direction of higher voltage (higher potential); when travelling in the opposite direction, it will be retarded. In an electric field all points having the same voltage are conceived as being connected by so-called equipotential lines. If an electron passes obliquely through an electric field (e.g., between two electrically charged wire grids, Fig. 2), it will undergo an additional acceleration towards the lines of higher potential: it changes its direction of motion, i.e., it is "refracted". On the basis of this principle, electrostatic "lenses" for electrons were constructed—in analogy with glass lenses for light—from spherically curved grids of wire netting (Fig. 3). However, electronic lenses of this kind have disadvantages, and for this reason the lenses now employed are based on tubes (Fig. 4) or diaphragms with apertures (Fig. 5). The fact that, although the equipotential lines present a symmetrical pattern, the concentrating and the dispersing action do not cancel each other out is because the electrons traverse the dispersing region at higher velocity, so that they undergo less deflection. Besides electrostatic lenses, there are also magnetic lenses. The function of a magnetic lens is based on the following principle: A moving electron is the most elementary form of an electric current; it is therefore surrounded by a magnetic field. When the magnetic field associated with an electric current interacts with another magnetic field, forces are exerted on the conductor through which the current flows. This is the basic principle of all electric motors and generators. For the same reason an electron travelling in a magnetic field undergoes a change of direction. To obtain high field strengths, magnetic lenses are made of iron-encased coils provided with a narrow gap (Fig. 6).

In principle the construction of an electron microscope—with electrostatic or with magnetic lenses—is very similar to that of a light microscope for photographic recordings (Figs. 7a, 7b, 7c): The electrons are emitted by the incandescent cathode, accelerated, and concentrated by the condenser on to the object to be examined. The object—a bacterium, a virus or a so-called replica—is supported on an extremely thin collodion film. (A replica is an envelope of carbon or some other suitable substance formed on the surface of a metal or mineral and then removed for observation under the electron microscope). Depending on the thickness and composition of the object, the electron rays undergo varying degrees of attenuation. The objective lens forms them into the enlarged intermediate image. This in turn is used to form a further enlarged image by the projector lens system. This image may be projected on to a fluorescent screen (for visual observation) or on to a photographic plate sensitive to electron rays.

Green light has a wavelength of around 1/2000 millimetre. Electrons accelerated with 50,000 volts have a 100,000 times smaller wavelength. Because of this extremely short wavelength, the resolving power of an electron microscope is very much greater than that of a light microscope. In combination with further enlargement of the photographic image, magnifications ranging from 100,000x to 500,000x can be obtained.

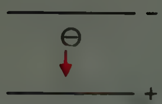


Fig. 1 ELECTRON IN AN ELECTRIC FIELD

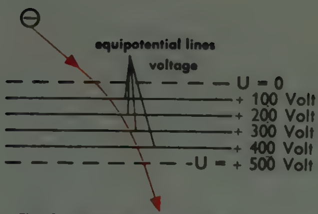


Fig. 2 "REFRACTION" OF THE ELECTRON BEAM

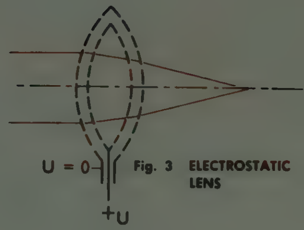


Fig. 3 ELECTROSTATIC LENS

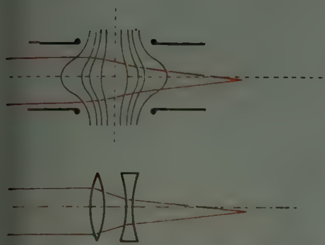


Fig. 4 TUBULAR ELECTRONIC LENS COMPARED WITH A CORRESPONDING SYSTEM OF OPTICAL LENSES

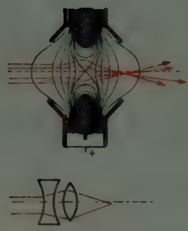


Fig. 5 DIAPHRAGM TYPE ELECTRONIC LENS COMPARED WITH A CORRESPONDING SYSTEM OF OPTICAL LENSES

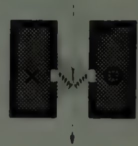


Fig. 6 IRON-ENCASED COIL AS A MAGNETIC LENS

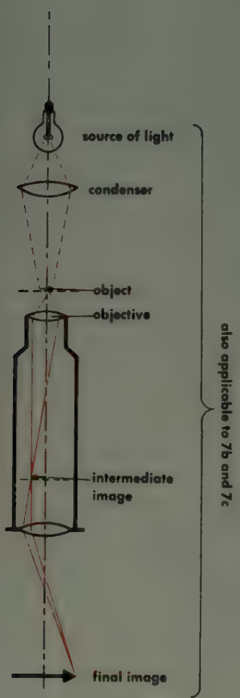


Fig. 7a OPTICAL MICROSCOPE

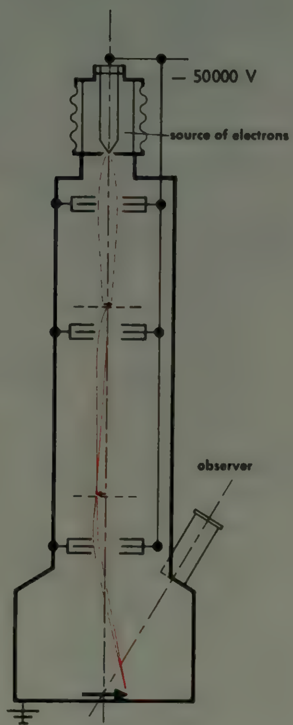


Fig. 7b ELECTROSTATIC ELECTRON MICROSCOPE (diaphragm type lens)

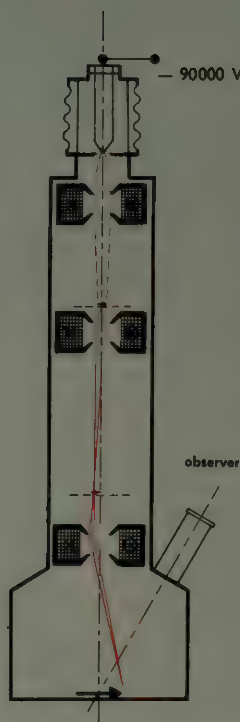


Fig. 7c MAGNETIC ELECTRON MICROSCOPE (coil type lens)

Microscope accelerator for U.S. Steel Corporation's million-volt electron microscope
Photo USIS

Telescopes are often said to "enlarge things" or "bring distant objects nearer". Actually they increase the visual angle. We are accustomed to judging the size or distance of objects by the angle under which we see them (Fig. 1).

In principle, all refracting telescopes (as distinct from reflecting telescopes) comprise an *objective* (object glass), which is directed towards the object to be observed, and an *ocular* (eyepiece), to which the observer applies his eye. The rays coming from the distant object are almost parallel; they converge to form an image at the focus of the objective (cf. lenses, page 158). This point also coincides with the focus of the ocular, so that the rays emerging from the latter are again parallel. The observer thus sees the object as though it were at infinity, but under a larger angle than without the aid of the telescope. The magnification is defined as the ratio of the focal length of the objective (f_1) to that of the ocular (f_2), i.e., the ratio f_1/f_2.

In the so-called Galilean telescope (named after Galileo, the great Italian scientist and astronomer) (Fig. 2) the ocular is a divergent lens. This system produces an upright image. This arrangement is nowadays used more particularly in opera-glasses, which are low-powered binoculars (usually with a magnification of $2\frac{1}{2}\times$).

The so-called *Keplerian telescope* (named after the German astronomer Kepler) (Fig. 3) has a convergent lens for its ocular. The fact that the intermediate image in this type of instrument is a real image is a great advantage: everything located in the plane of this image appears sharp and as though it were at infinity. For this reason a diaphragm is installed in this plane, so as to form a sharp boundary to the field of vision. Also, crosswires or a glass measuring scale or some other device—depending on the purpose for which the telescope is used—may be installed here. The inverted position of the image is no objection when the telescope is used for astronomical observations. For terrestrial use, however, it is necessary to introduce an extra lens to produce a second intermediate image, this time in the upright position. For this purpose the "terrestrial telescope" is provided with an additional convergent lens between the objective and the ocular (Fig. 4). This intermediate lens does not enlarge the image formed by the objective, but merely "brings it the right way up". The drawback of this type of telescope is that, because of the intermediate inversion of the image, the instrument has a rather large overall length. For this reason the tube is often of collapsible construction, consisting of segments which can be slid into one another ("telescoped"). The old portable telescope constructed on this principle has now largely been superseded by prismatic binoculars (page 174).

It is impracticable to make lenses of more than about 40 in. diameter. For this reason, in very large astronomical telescopes a concave mirror instead of a lens is used as the objective (the big mirror of the world's largest telescope, at Mount Palomar in the U.S.A., has a diameter of 200 in.) Fig. 5 shows the principle of the reflecting telescope, and Fig. 6 gives an example of a combination of a concave main mirror and a convex collecting mirror; the overall length of the instrument is quite small.

In astronomical telescopes used for photographic purposes the photographic plate is placed direct in the image plane of the objective.

The principle of the reflecting telescope with its conveniently short construction length has also been introduced into ordinary photography. Fig. 7 shows the optical arrangement of a miniature camera comprising a mirror objective with a focal length of 20 in. and an aperture of $f/4.5$. It consists of two glass mirrors and four low-powered correcting lenses.

Fig. 1 INCREASE OF THE VISUAL ANGLE

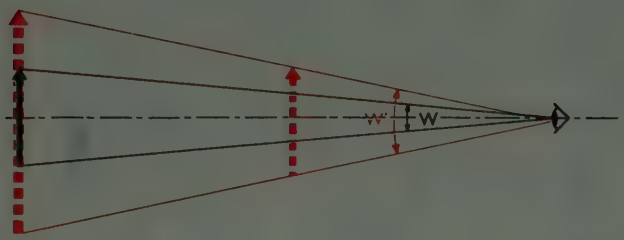


Fig. 2 GALILEAN TELESCOPE

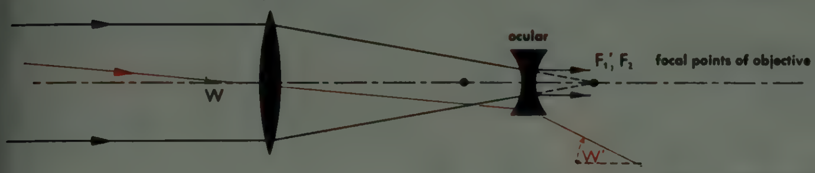


Fig. 3 KEPLERIAN TELESCOPE

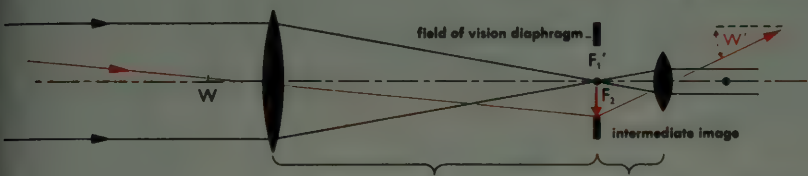


Fig. 4 TERRESTRIAL TELESCOPE

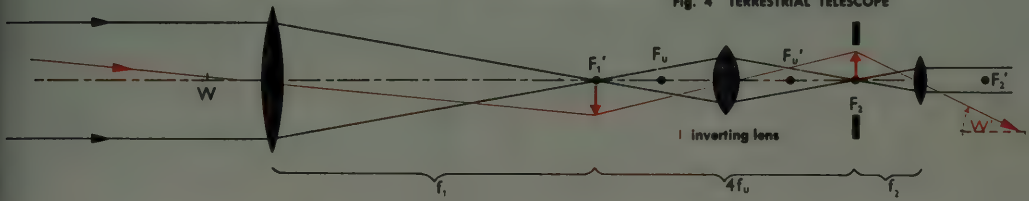
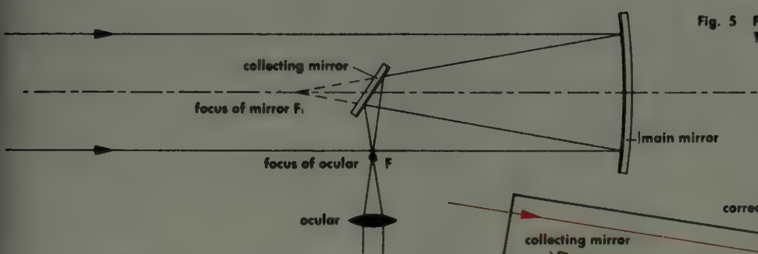


Fig. 5 PRINCIPLE OF THE REFLECTING TELESCOPE



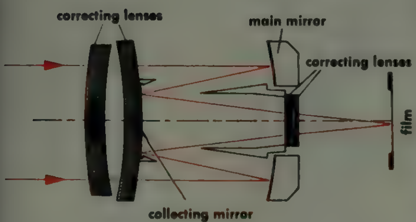


Fig. 7 MIRROR LENS, F/4.5, 500 MM FOCAL LENGTH, FOR MINIATURE CAMERA (*24 mm x 36 mm picture size*)

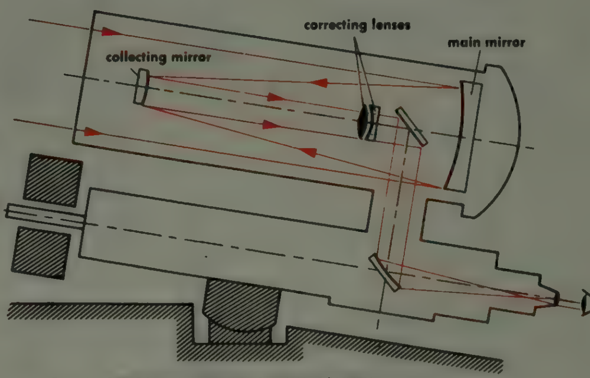


Fig. 6 REFLECTING TELESCOPE WITH CONVEX COLLECTING MIRROR AND CORRECTING LENSES

The 200-inch mirror tilted for testing at the California Institute of Technology
Photo USIS

Prismatic binoculars (or, in its single form, the prismatic telescope) comprises an objective and an ocular, like the astronomical telescope (see page 170). The latter, however, produces an inverted image. In the prismatic telescope the image is brought into the upright position by means of prisms. Fig. 1. shows how a beam of light is twice deflected at right angles in a prism (by total reflection at the boundary surfaces), with the result that "top" and "bottom" are interchanged. However, in order to rectify the inverted image formed by the objective, it is necessary also to interchange "left" and "right". For this reason the light is passed through a second reversing prism (Fig. 2). In comparison with the terrestrial telescope (page 170) there is a very substantial saving in the overall length of the instrument. Besides, the two object lenses are spaced farther apart than the eyepieces; this makes for better stereoscopic seeing. Fig. 4 shows the course of rays in prismatic binoculars. The point of convergence of the rays in front of the ocular marks the position of the intermediate image. Here, too, is located the diaphragm forming the boundary of the field vision.

The magnification of a telescope is usually indicated in combination with the diameter of the objective. For instance, 8 × 30 means the magnification is 8 × (eight times) and that the objective has a diameter of 30 mm. The ratio of objective diameter to magnification—i.e., in this case 30/8 = 3.75—is an important criterion. It means that the beam of rays emerging from the so-called exit pupil of the instrument has a diameter of 3.75 mm. The exit pupil is the image of the eight times larger entrance pupil. These two pupils can be seen on looking into the ocular or the objective from a distance of about 12 inches. The amount of light entering the human eye, too, is controlled by the pupil; in feeble light its diameter may be as much as 8 mm, but in bright sun it diminishes to 1.5 mm. The light-gathering power of a telescope or field-glass is therefore best utilised when the beam of rays emerging from the ocular has the same diameter of the pupil of the observer's eye. So-called "night-glasses" are characterised, for example, by the data 7 × 50, i.e., the magnification is 7 × and the objective diameter (= entrance pupil) is 50 mm. In moonlight the entire emergent beam enters the observer's eye (Fig. 3a). On the other hand, if the instrument is used in bright sunlight, the pupil of the eye is, say, 2 mm in diameter (Fig. 3b), so that then only $\frac{1}{13}$ of the instrument's light-gathering power is utilised. As the amount of light passing through the aperture is proportional to the area thereof, the light-gathering power of a telescope is designated by the square of the diameter of the exit pupil, e.g., 7.14 × 7.14 = 51.5.

In twilight, objects are distinguished better according as they appear larger and brighter. It is known that the perceptibility increases approximately in proportion to the square root of the product of objective diameter and magnification, e.g.,

$$\sqrt{30 \times 8} = 15.5; \qquad \sqrt{50 \times 7} = 18.7.$$

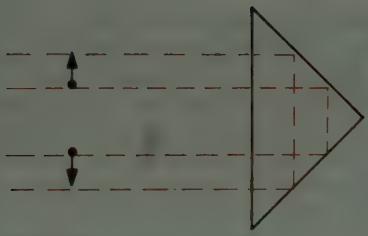


Fig. 1 TOTAL REFLECTION AT TWO FACES OF A PRISM

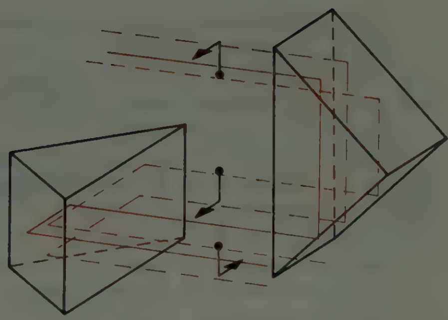


Fig. 2 REVERSAL OF IMAGE BY TWO PRISMS

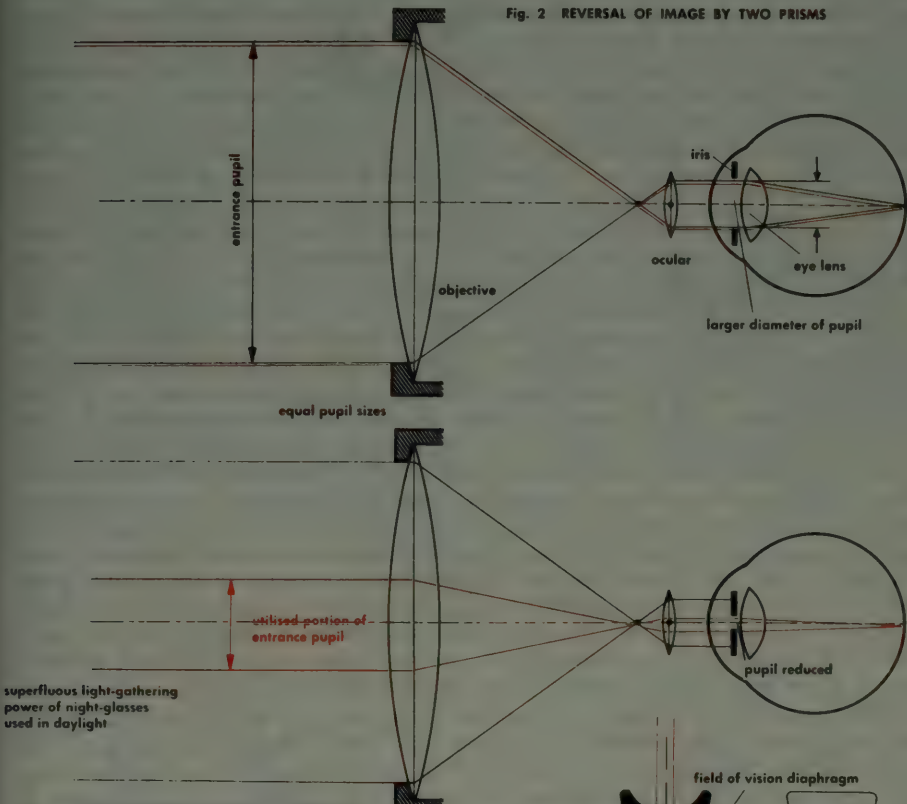


Fig. 3 LIGHT-GATHERING POWER OF TELESCOPE AND EYE (e.g., NIGHT-GLASSES)

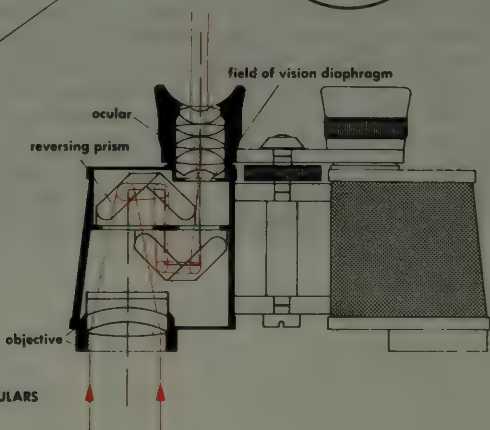


Fig. 4 PRISMATIC BINOCULARS

A camera is a device consisting essentially of a light-tight box which has an opening covered by a lens in one wall; this lens forms a real image of the object upon a plate or film of light-sensitive material disposed inside the box. A modern camera consists of a body, or housing, with the film holder and the film feed mechanism, the lens, the shutter, the distance setting (possibly also a rangefinder), and the viewfinder for composing the picture. A small portion of the light from the object is momentarily allowed to form a real image on the photographic plate or film. The simplest method of forming an image of this kind is by means of a "pinhole" diaphragm, containing a hole of 0.016 in. diameter, on the front of the light-tight box. On this principle it is possible to construct a so-called pinhole camera (Fig. 1). From each point of the object a ray of light passes through the hole and strikes the photographic plate, where it makes an image of that point. Each point of the object is, admittedly, projected on to the plate as a small elliptical spot of light, and the adjacent ellipses merge together, so that a blurred image is formed. Besides, a pinhole camera has a low light intensity, i.e., it has a very low-speed "lens" (the pinhole), which means that long exposure times are necessary for obtaining adequate blackening of the plate. Sharper images are obtained by using an objective consisting of a lens or, preferably, a set of lenses (elements) mounted one behind the other. In that case the light coming from each point of the object to be photographed is focused by the lens and concentrated as a corresponding point on the photographic plate. In a box camera (Fig. 2) the aperture and therefore the light-transmitting power of the lens are fixed. This type of camera usually has a single or a two-element lens. In more advanced cameras both the aperture and the exposure time can be varied to suit the lighting conditions and/or the movement (if any) of the object (moving objects must be photographed with short exposures in order to obtain sharp images on the plate or film). The lenses in these more elaborate cameras usually consist of three or more elements. Cameras using the larger sizes of film are usually constructed as bellows-type folding cameras (Fig. 3). Smaller cameras are often of the tube type (Fig. 4). This latter form of construction, as applied to so-called miniature cameras, has important optical advantages (short focal length, larger angular field and the resultant advantages. High-class models are equipped with coupled rangefinders, interchangeable lenses, and shutters with speeds up to $\frac{1}{1250}$ second. The twin-lens reflex camera (Fig. 5) comprises a camera part and a viewfinder part. The distance setting for the two parts is coupled, so that the sharply focused image which is formed on the ground-glass screen of the viewfinder is also formed on the film plane. The lenses of the two parts have exactly the same focal length. As a rule, the viewfinder lens has greater light-transmitting power than the camera lens, as this makes for quicker and sharper focusing. All twin-lens reflex cameras take $2\frac{1}{4}$ in. $\times$ $2\frac{1}{4}$ in. photographs. The single-lens reflex camera (Fig. 6) has only one lens. This lens first serves for focusing the image on a ground-glass screen (or in a prism viewfinder); then, when the image has been focused, the deviating mirror (set at 45°) is swung upwards, enabling the image to be formed on the film plane. The stop value (aperture) is set before the exposure is made. When the shutter is released, the deviating mirror is automatically swung out of the way, and the shutter, which in such cameras is usually of the focal plane type, i.e., mounted directly in front of the film (see page 186), momentarily exposes the film. Then, when the film is wound on for the following shot and the shutter is thereby automatically cocked, the mirror is lowered again for focusing the next picture.

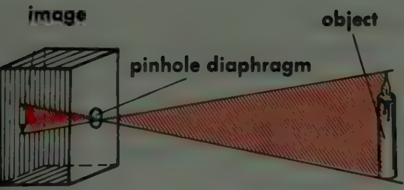


Fig. 1 PINHOLE CAMERA

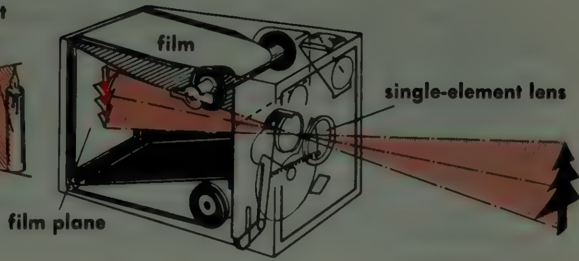


Fig. 2 BOX CAMERA

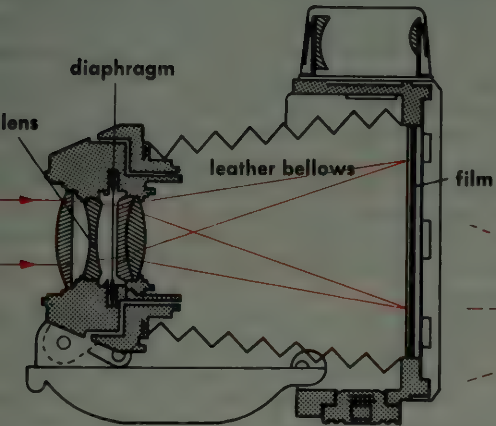


Fig. 3 BELLOWS-TYPE FOLDING CAMERA

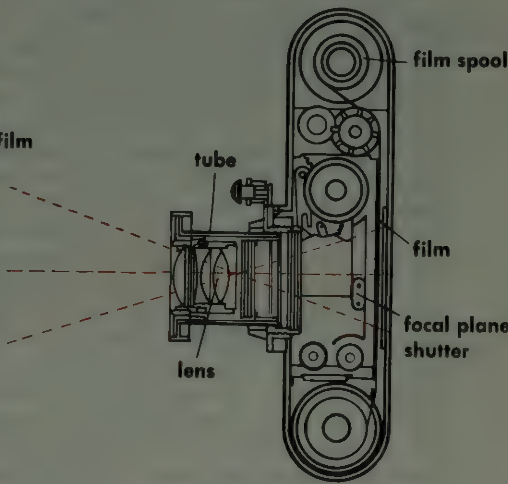


Fig. 4 TUBE-TYPE CAMERA

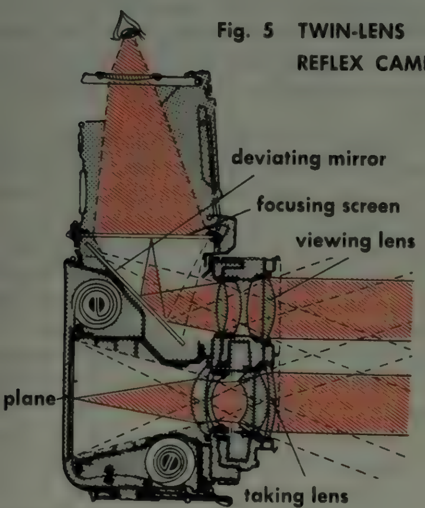


Fig. 5 TWIN-LENS REFLEX CAMERA

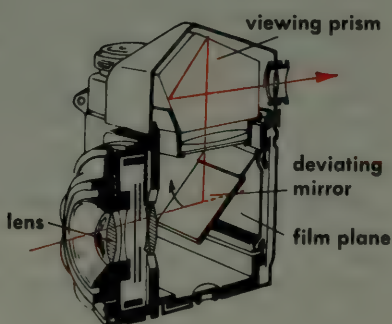


Fig. 6 SINGLE-LENS REFLEX CAMERA

The most important component of a camera is the objective or, as it is more familiarly called, the lens. The lens should form a sharp, even, undistorted and bright image over the entire picture size.

The simplest conceivable "lens" is merely a small hole—a "pinhole"—which forms an image on the back wall of the camera (where the photographic plate or film is mounted) because light travels in a straight path (Fig. 1). However, a photograph taken with a pinhole camera is not very sharp, as a point is reproduced in the image as a spot of light which can never be smaller than the pinhole itself. Also, because of the very small size of the hole, its light-transmitting power is extremely low (see page 188).

Because of its ability to concentrate the rays in a beam of light at one point, a lens is the obviously suitable device for forming a suitably sharp and bright image in a camera. Fig. 2. shows diagrammatically how the image is formed by the lens (see page 158). However, rays entering the lens at wide angles in relation to the optical axis are not ideally concentrated at one point: the deviations from the ideal image are called lens errors (or lens aberrations). The principal errors are: spherical aberration (Fig. 3; the marginal rays are focused at a point closer to the lens), astigmatism and curvature of the image field (Fig. 4; two mutually perpendicular rays passing obliquely through a lens are concentrated on two curved surfaces of different curvature), coma (comet-like elongation of points outside the centre of the image), distortion (Fig. 5), chromatic aberration (Fig. 6; light rays of different colours are not focused at the same point: light of longer wavelength is less strongly refracted than shorter-wave light; see page 156).

If a convergent lens is combined with a suitable divergent lens, which is so shaped as to have half the (negative) refractive power of the convergent lens but is made of glass with twice the colour dispersion, the refractive power of the lens combination is halved, but the chromatic aberration is entirely eliminated (Fig. 7). A lens (or, strictly speaking, an objective comprising two elements) of this kind is called an achromatic lens. Its two component elements are usually cemented together (Fig. 8), and it is used for box cameras with apertures up to $f/9$. The other lens errors can be corrected by similar means, i.e., by the combination of several elements made of different kinds of glass possessing different optical properties. Hundreds of varieties of glass differing in refractive index and dispersion are available to the lens designer. By varying the number of lenses ("elements"), the type of glass, the radii of curvature, the lens thicknesses, and the air gaps (if any) between the elements, it is possible to reduce the optical errors to acceptable values.

Besides the other errors, anastigmatism is a defect which is particularly necessary to eliminate in the case of a "fast" lens, i.e., a lens with a high light-transmitting power. A lens corrected in this way is called an anastigmatic lens (or "anastigmat"). Some fundamental forms of lens in this category have proved satisfactory. Thus, the so-called "*triplet*" (Fig. 9) is fitted as the standard lens in nearly all cameras in the medium price range, with apertures of $f/3.5$ or $f/2.8$ (see page 188) and focal lengths of 45 to 50 mm. These lenses are known by various proprietary names (e.g., Agnar, Apotar, Cassar, Lanthar, Novar, Pantar, Radionar, Reomar, Triotar). An important form of camera lens is the cemented four-element triplet (Fig. 10), which usually has an aperture $f/2.8$ and a focal length of 50 mm; for larger sizes of camera these lenses have an aperture of $f/3.5$ for a focal length of 75 mm or $f/4.5$ for a focal length in 105–300 mm range (e.g., Elmar, Skopar, Solinar, Tessar, Xenar, Ysarex). Also, there are five-element variants (e.g., Apo-Lanthar, Elmarit, Heliar), and four- to seven-element variants of the Sonnar lens, which can hardly be classed as triplets at all. For large angular fields, symmetrical arrangements of the lens elements facilitate the correction of the image defects. Most of the fast lenses used in miniature cameras, with apertures of about $f/2$, are of the *Gauss double anastigmat* type (Fig. 11) (e.g., Biotar, Heligon, Pancolar, some Planar lenses, Auto-Quinon, Solagon, Ultron, Xenon). In some lenses a seventh or indeed an eighth element is added in order to obtain good correction even at the highest light-transmitting power ($f/2$ to $f/1.4$) (Nokton, some Planar lenses, Septon, Summicron, Summilux). For further combinations see the article on interchangeable lenses (page 182).

Fig. 1 PINHOLE CAMERA

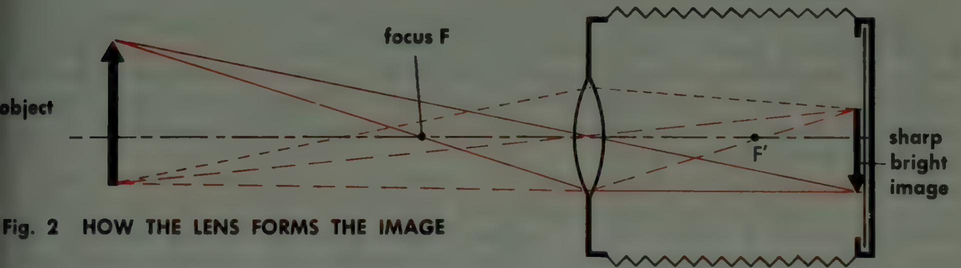


Fig. 2 HOW THE LENS FORMS THE IMAGE

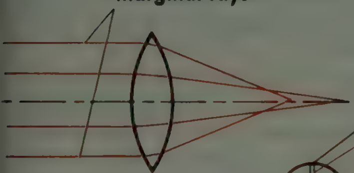


Fig. 3
SPHERICAL ABERRATION

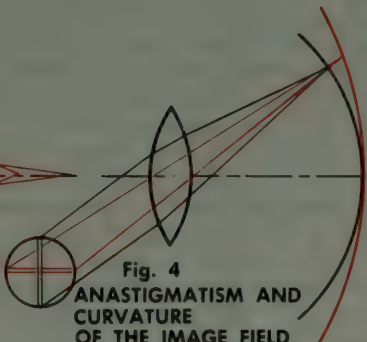


Fig. 4
ANASTIGMATISM AND CURVATURE OF THE IMAGE FIELD

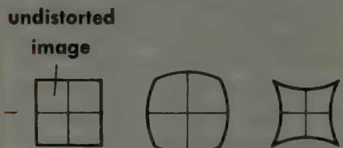


Fig. 5 DISTORTIONS

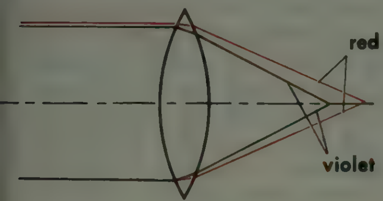


Fig. 6 CHROMATIC ABERRATION

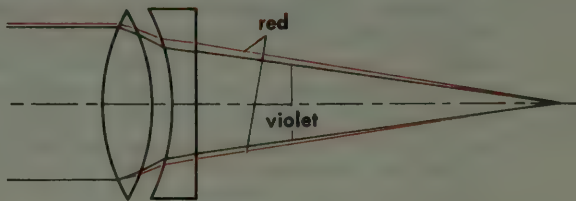


Fig. 7 CORRECTION OF CHROMATIC ABERRATION

Fig. 8
ACHROMATIC LENS

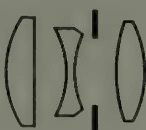


Fig. 9
ANASTIGMATIC LENS (*triplet*)

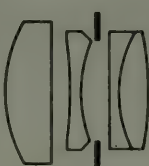


Fig. 10
CEMENTED TRIPLET

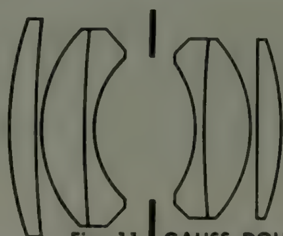


Fig. 11 GAUSS DOUBLE ANASTIGMAT

basic forms of photographic lenses (in each case the object to be photographed is assumed to be situated to the left of the lens)

The size of the image in a camera is determined by the distance between the lens and the film plane.

Fig. 1a shows the conditions in a camera provided with a "standard lens", while Fig. 1b is a diagram of a camera with a "wide-angle lens". The latter has a wider angular field, but the actual image size is the same (for the same film size). From Fig. 1c we see: the image size B bears the same ratio to the subject size G as does the image distance b to the subject distance g:

$$\frac{B}{G} = \frac{b}{g} \quad \text{or alternatively:} \quad \frac{B}{b} = \frac{G}{g}$$

For large "taking distances" (distance from camera to subject) the image distance b is approximately equal to the focal length f (cf. the article on lenses, page 158); hence the image size is, generally speaking, proportional to the focal length: a short focal length produces a small image of the subject; if the focal length is, say, three times as large, the image will also be three times as large (Figs. 3a and 3b; principle of wide-angle and telephoto photographs).

The image size must now be considered in relation to the picture size. The focal length of a "standard lens" is about equal to the diagonal of the picture size, i.e., a miniature camera or a plate camera with standard lenses both form an image of a section of space whose diagonal is therefore approximately equal to the distance from the camera. This corresponds to an angular field of 53° over the diagonal. The usual standard lenses have an angular field in the range of 45° to 60°, corresponding to focal lengths of 45–50 mm for miniature cameras with picture size 24 mm × 36 mm (diagonal 43 mm), 75–85 mm for picture size 6 cm × 6 cm (diagonal 85 mm), and 135–150 mm for picture size g × 12 cm (diagonal 150 mm).

The question as to the angular field for which a lens system is corrected is of significance only with regard to cameras having extension bellows, because in such cameras a lens of the same focal length is used as a telephoto lens for a roll film size, as a standard lens with reserve scope for adjustment, or as a wide-angle lens for a larger plate size, provided that the camera is suitably designed for such a large angular field (Fig. 2). In a miniature camera the lens is normally mounted in a tube of appropriate length and is corrected exactly for the angle of the particular picture size employed (see the article on interchangeable lenses (page 182).

If different focal lengths are used at one and the same view point, the image size will be altered, but not the perspective effect: in that case a photograph taken with a telephoto lens will be no different from a wide-angle photograph which is subsequently enlarged (Figs. 3a and 3b). A much more favourable method is to use different focal lengths to alter the perspective by changing the distance. Figs. 3b and 3c show how the size ratio between foreground and background is reversed: the distance to the foreground subject V is adapted to the focal length so that the image V^1 retains its size; the distance between the foreground and the background does not change, however, so that here the change affects the scale of the image.

Fig. 1 DIFFERENT FOCAL LENGTHS, SAME IMAGE SIZE

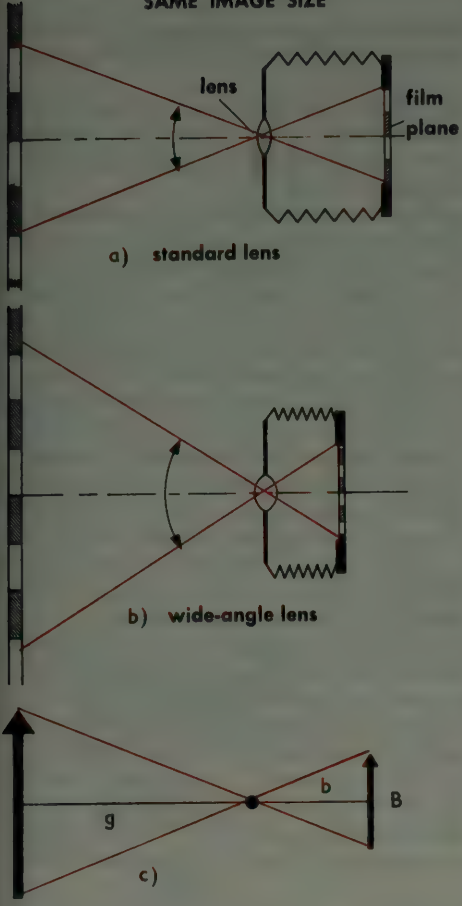
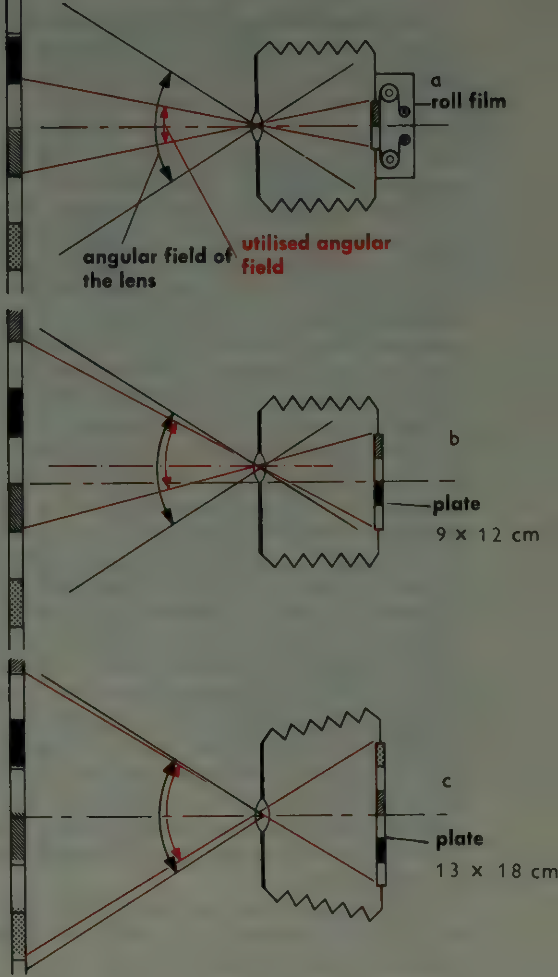


Fig. 2 SAME FOCAL LENGTH, DIFFERENT IMAGE SIZES



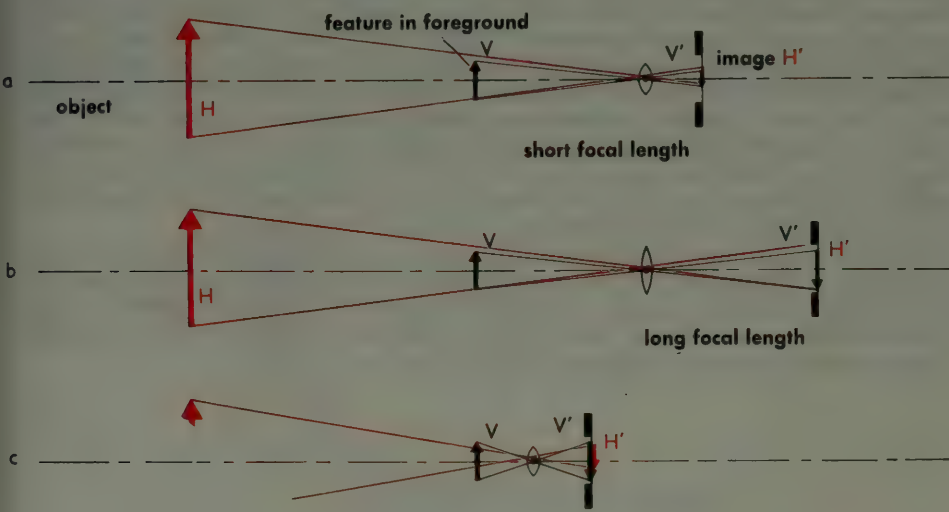


Fig. 3 FOCAL LENGTH AND PERSPECTIVE

Some cameras are so constructed that the focal length can be changed, in order to enable the scale of the picture (or image) to be suited to the size of the subject and to modify the perspective (cf. page 180).

The systems for changing the focal length and the construction of these interchangeable lenses will here be explained with reference to the three focal lengths most commonly employed for miniature cameras (wide angle lens approx. 35 mm, standard lens approx. 50 mm, portrait lens approx. 85 mm).

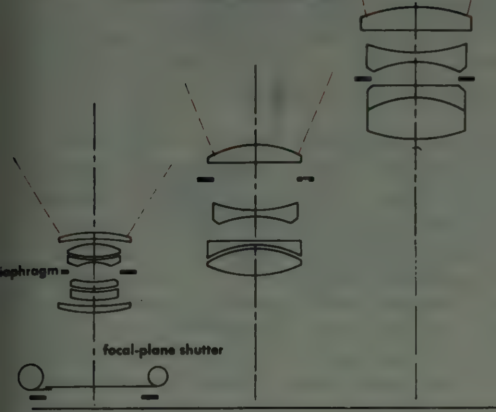
Unlimited possibility of changing the lenses is afforded by the focal plane shutter (cf. page 186, Fig. 1). In this example the standard lens (1b) is a triplet with a cemented rear element. Because of its shorter focal length, the wide angle lens (Fig. 1a) is mounted closer to the film plane, whereas the telephoto lens (1c), with its greater focal length, has a correspondingly longer tube. For cameras fitted with focal plane shutters there are lenses with focal lengths ranging from 21 mm to 600 mm. Indeed, there are special lenses with focal lengths of as little as 8 mm or as much as 2000 mm, corresponding to a scale ranging from 0.16 to 40 times in relation to the standard lens.

Single-lens reflex cameras require a certain amount of space for the collapsible mirror behind the lens, and this "width of cut" represents a minimum distance between the back of the lens and the image. Similar limitations occur in the case of a camera with a diaphragm shutter (cf. page 184) if it is designed for interchangeable lenses: instead of being mounted between the lenses, the shutter must then be installed directly behind the interchangeable lens, so as to ensure that even with very short exposures there is not too much lack of sharpness at the corners of the picture. If this requirement is fulfilled for the standard lens (2b), then the width of cut, i.e., the distance between the back of the lens and the image, will be predetermined, just as in the mirror reflex camera. However at some additional cost it is nevertheless possible to make a wide angle lens whose focal length is shorter than the width of cut (Fig. 2a). Greatly simplified, this lens could be described as a standard lens in which the path of the rays is modified by the divergent element. The arrangement illustrated in (2c) is a "true telephoto lens" in that it comprises a "positive" (convergent) and a "negative" (divergent) set of lenses; together these form a system having a large focal length, in the manner of a Galilean telescope (see page 170). This form of construction has the advantage of requiring only a relatively short construction length (cf. Figs 2c and 1c).

A third possibility is the so-called convertible lens. The lens with shutter and diaphragm is permanently connected to the body of the camera (just as in a camera without interchangeable lenses), only the front element of the lens being interchanged. The combinations illustrated in Figs. 3a and 3c correspond approximately to the wide angle lens and the telephoto lens of Figs. 2a and 2c respectively. It is on this principle, too, that the functioning of the so-called "zoom lens", which allows of "infinitely variable" control of focal length, is based. Fig. 4 shows a lens of this kind in the "telephoto" setting. The front set of lenses as a whole is positive, whereas the rear set is negative. By and large, the lens resembles a telephoto lens as shown in Fig. 2c. Now if the two movable sets of lenses are slid backwards (towards the camera) a continuous transition is effected until an arrangement rather like that in Fig. 2a is reached, i.e., the lens now functions as a wide angle lens.

Zoom lenses with large focal length ranges—e.g., 8 to 48 mm for 8 mm cine film—are now extensively used on television cameras and narrow-gauge cine cameras. For the latter class of cameras the cost of such lenses is not prohibitively high (because of the small picture size), and attractive or interesting effects in motion pictures can be obtained by means of zoomed shots.

Fig. 1 LEICA INTERCHANGEABLE LENSES FOR UNRESTRICTED APPLICATION

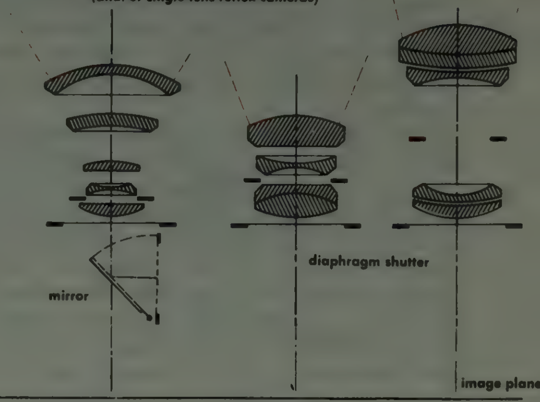


a) Summaron 1:2.8 f=35mm

b) Elmar 1:2.8 f=50mm

c) Elmar 1:4 f=90mm

Fig. 2 RODENSTOCK INTERCHANGEABLE LENSES FOR DIAPHRAGM SHUTTER
(and/or single-lens reflex cameras)

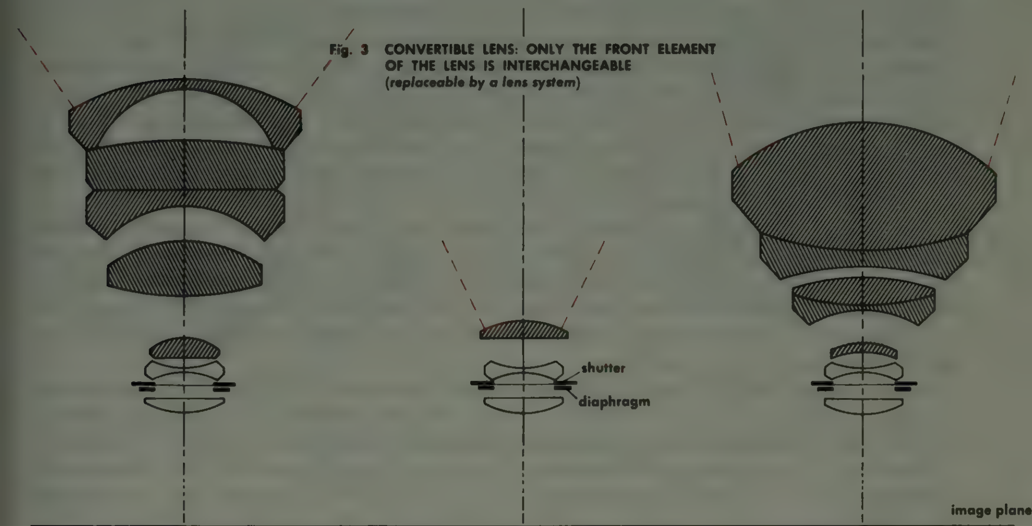


a) Eurygon 1:4 f=35mm

b) Ysarex 1:2.8 f=50mm

c) Rotelar 1:4 f=85mm

Fig. 3 CONVERTIBLE LENS: ONLY THE FRONT ELEMENT OF THE LENS IS INTERCHANGEABLE
(replaceable by a lens system)

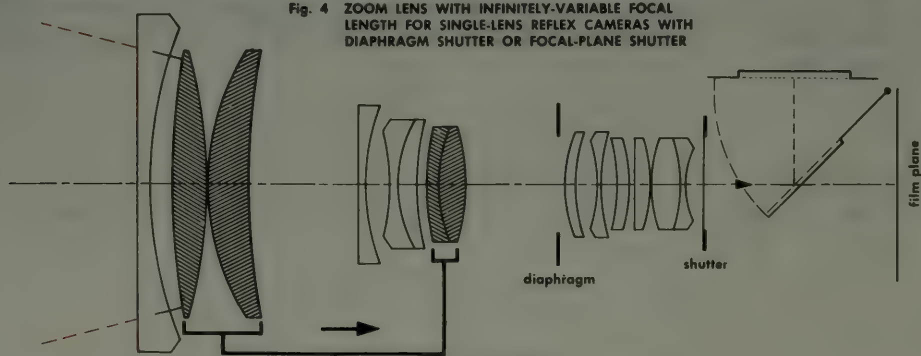


a) Pantar in wide-angle combination 1:4; f=30mm

b) Pantar 1:2.8; f=45mm
Zeiss

c) Pantar in telecombination 1:4; f=75mm

Fig. 4 ZOOM LENS WITH INFINITELY-VARIABLE FOCAL LENGTH FOR SINGLE-LENS REFLEX CAMERAS WITH DIAPHRAGM SHUTTER OR FOCAL-PLANE SHUTTER



Zoomar (Voigtländer) 1:2.8; f = 36 to 82 mm
here shown in telephoto setting (82 mm); by displacement of the two shaded elements the focal length can be continuously shortened

The function of a camera shutter is to admit the light rays to the film for a controlled—and usually very brief—length of time. Diaphragm shutters are usually in the form of a "between-the-lens" shutter, i.e., the shutter is mounted in the diaphragm plane between the individual elements which the compound lens of the camera is composed. In particular cases, however, the shutter may be located directly behind the lens.

This type of shutter comprises a number of thin steel plates (usually five), known as shutter leaves or shutter blades, which rotate about pivots in such a way that they open out in a direction away from the optical axis, i.e., from the centre of the lens outwards, and then, on completion of the exposure, close the lens by moving back towards the centre. Before the exposure is made, the shutter has to be cocked by tensioning a spring ("pre-set shutter"). This cocking operation is often coupled with the film advance motion. Simple "everset" shutters (with speeds of approx. $\frac{1}{30}$ to $\frac{1}{125}$ sec.) are cocked only when the releaser is depressed. The functioning of the shutter is shown simplified, with reference only to one shutter leaf instead of the actual five leaves, in Figs. 1a to 1d. In Fig. 1a the shutter is still closed. Now the drive element rotates in the arrowed direction (Fig. 1b) and thrusts against the shutter-opening pin of the actuating ring, causing the latter to rotate anti-clockwise. Other pins mounted on this ring engage with slots in the shutter leaves and cause these to swing about their respective pivots. In Fig. 1c the shutter is fully opened. Further rotation of the drive element is now prevented by an escapement mechanism, which is really a kind of clockwork in which an oscillating anchor controls the movements of the cog-wheels. When the exposure period (up to 1 sec.) has ended and the escapement has run down, the drive element continues its rotation in the same direction as before. A projection on the drive element now thrusts against the shutter-closing pin on the actuating ring and causes this ring to rotate in the reverse direction, so that the shutter leaves close again.

The actual opening and closing movements are accomplished in about 2 milliseconds ($=\frac{2}{1000}$ sec., see Fig. 2a). The "shutter open" time regulated by the escapement is therefore about 2 milliseconds shorter than the nominal exposure time to which the shutter mechanism has been set. For very short exposures ($\frac{1}{50}$ sec. and less; Fig. 2b) the escapement is disconnected and an auxiliary spring speeds up the shutter movements still more.

Some cameras are equipped with a self-timer (automatic release). When this device is used, actuation of the releaser first sets a delayed-action mechanism in motion which initiates the movement of the shutter leaves after an interval of 5 to 10 seconds.

For flashlight photography the actual flash must be synchronised with the opening of the shutter. Most shutters are provided with a so-called *X* contact which establishes electrical contact at the instant when the shutter is fully opened (electronic flash). "Fully synchronised" shutters are additionally equipped with an *M* contact, which ignites the flash about 16 milliseconds earlier, so as to give the somewhat slower flash produced by a flashbulb sufficient time to attain its maximum intensity (cf. page 208).

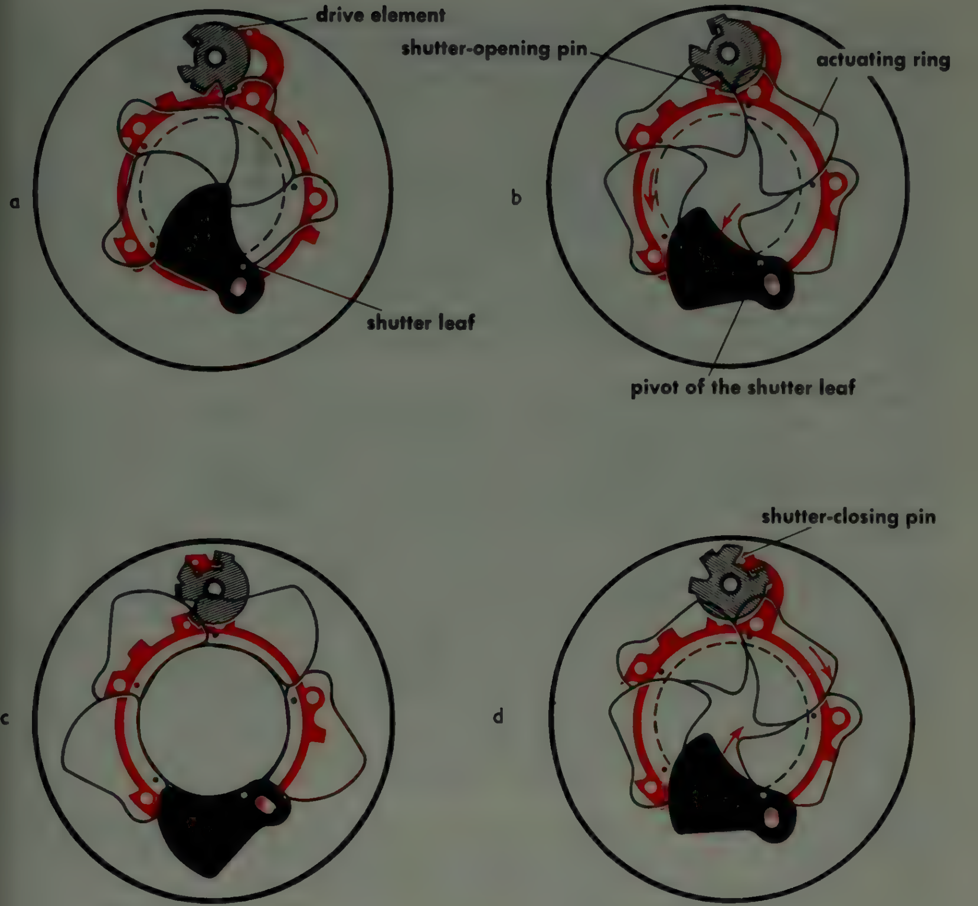


Fig. 1 SEQUENCE OF MOVEMENTS OF A SHUTTER LEAF

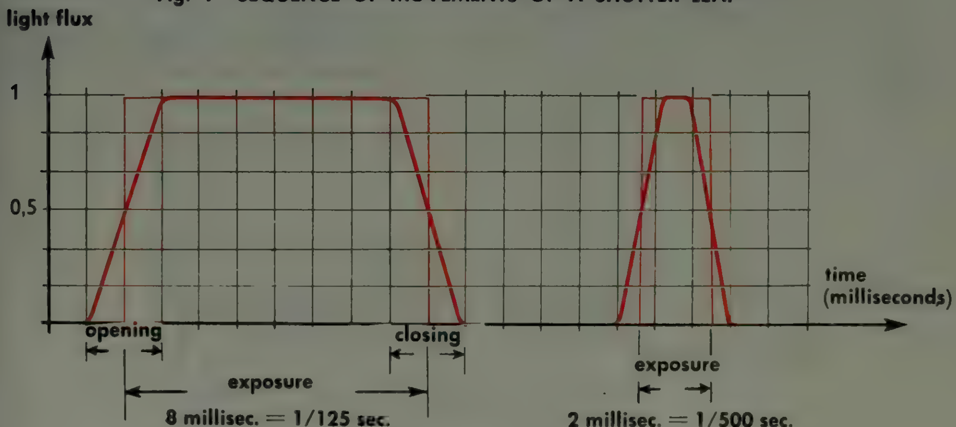


Fig. 2 LIGHT FLUX / TIME DIAGRAM OF A DIAPHRAGM SHUTTER

If the camera is fitted with interchangeable lenses for long and short focal lengths respectively, or if the camera is to be used with close-up adapter rings and bellows attachment, or mounted on a microscope, then it must be equipped with a shutter which moves in a plane just in front of the film. A shutter of this kind is known as a focal plane shutter.

A focal plane shutter comprises two opaque roller blinds which move in the same guide tracks (Fig. 1). When the shutter release is actuated, the first blind is drawn aside and rolled on to a spool by means of a spring. After a certain pre-set interval of time the second blind unrolls at the same speed and covers the film again. The two blinds thus form a slit which travels across the film gate. If the shutter has been set to a long exposure (Fig. 2), the second roller blind will wait a relatively long time before following the first blind; in that case the slit is very wide. With a short exposure the second blind will quickly follow the first, so that only a very narrow moving slit is formed (Fig. 3).

The roller blinds may be designed to move either longitudinally across the picture area, i.e., a distance of 36 mm (as in Fig. 1), or transversely, i.e., a distance of 24 mm (as in Fig. 4). Most focal plane shutters of the roller blind type are designed to operate at speeds down to $\frac{1}{1000}$ sec. Such short exposures are obtained by narrowing down the slit, as already explained, while the speed of travel of the blinds remain the same. The entire duration of the exposure procedure is therefore substantially longer than $\frac{1}{1000}$ sec. If the subject of the photograph is moving rapidly past the camera, the image it forms on the film will be travelling either "along with" or "against" the direction of movement of the slit. This could cause a certain distortion (elongation or shortening) of the image, as was indeed found to occur in the older large cameras equipped with such shutters. However, in modern miniature cameras using 35 mm film the slit takes less than $\frac{1}{50}$ sec. to travel across the picture area. The image of a fast-moving object in front of the camera, such as a racing car, which takes, say, $\frac{1}{5}$ sec. to traverse the picture area would therefore undergo a distortion of 10%. The vehicle would have to be travelling very fast indeed for this effect to be objectionably noticeable.

The operating time of the shutter is, however, a very significant factor in flashlight photography. With short exposure times the brightness of the flashbulb must remain approximately constant for at least the length of time required for the slit to traverse the image. With an electronic flash or flashbulb with a very short flash period this condition would not be fulfilled, and in that case only a narrow strip of the film (corresponding to the width of the shutter slit at the actual time of the flash) would be exposed. To synchronise the camera to flashlight equipment of this kind it is therefore necessary so to adjust the exposure time that the slit width exceeds the width of the film gate at the instant of flash. In modern cameras this exposure time is about $\frac{1}{50}$ sec.

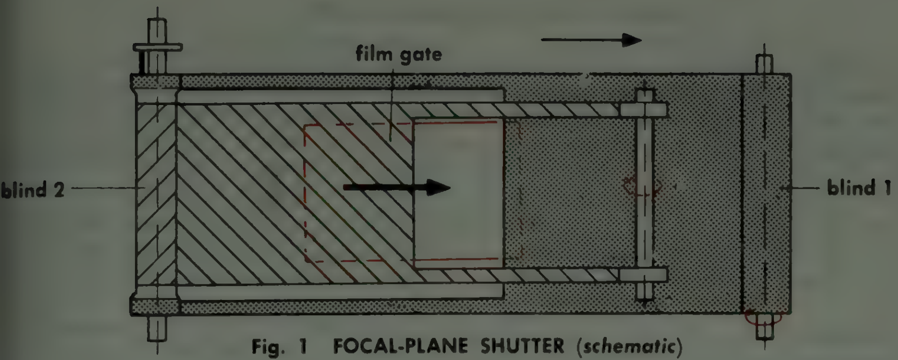
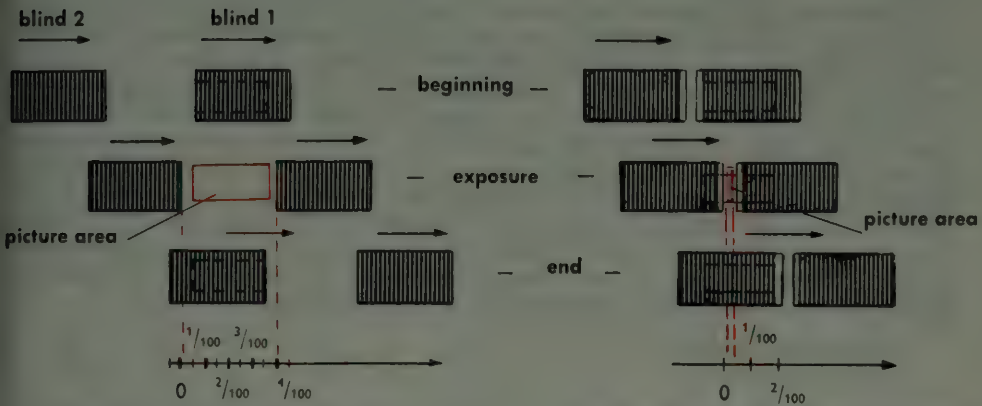


Fig. 1 FOCAL-PLANE SHUTTER (*schematic*)



g. 2 MOTION OF THE BLINDS FOR LONG EXPOSURE (1/25 sec.)

Fig. 3 MOTION OF THE BLINDS FOR SHORT EXPOSURE (1/200 sec.)

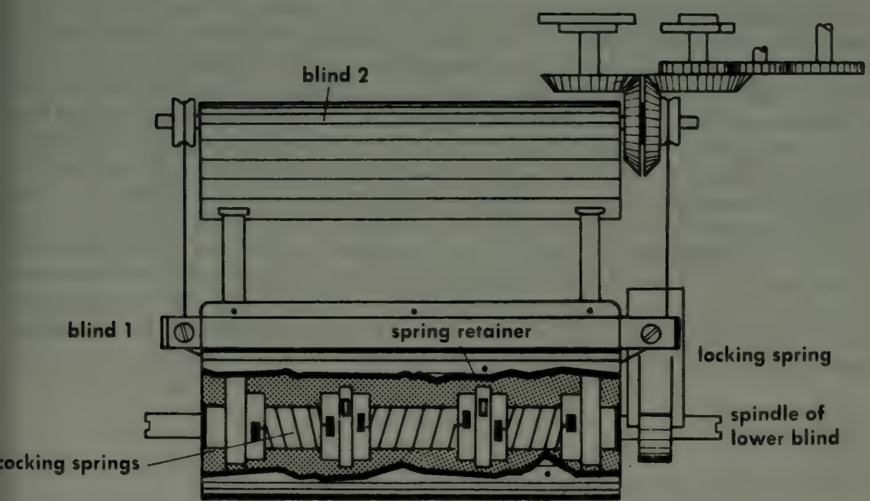


Fig. 4 MECHANISM OF A FOCAL-PLANE SHUTTER

SPEED OF A CAMERA LENS

The "speed" or light-transmitting power of a lens can be visualised by conceiving the camera as a closed chamber with an opening (a window) at one end (Fig. 1). The wall opposite the window will receive more light in proportion as the window is larger and the distance from the wall to the window is less. All the light rays entering the eye of an observer standing at the rear wall of the chamber will form a cone whose base is bounded by the edges of the window. The larger the window is, the wider will be the cone and the greater will be the amount of light entering the observer's eye. The ratio of the diameter of window opening (assuming it to be circular) to the length of the chamber is therefore a criterion of the "width" of the cone of light rays.

In a camera the rear wall is formed by the film, each point of which receives a quantity of light corresponding to the width (or diameter) of the aperture. Normally the distance between the lens and the film is approximately equal to the focal length of the lens. The larger the lens is, the more light it can transmit into the camera. For this reason the transmitting power of a lens is characterised by the ratio of the diameter of the effective aperture to the focal length. This is called the lens aperture ratio or relative aperture of the lens (Fig. 2). For example, if a camera lens has an aperture of 25 mm diameter and a focal length of 50 mm, its aperture ratio is 25:50 = 1:2, where the figure 2 is the so called "*f*-number" (often designated as *f*/2). If the aperture is reduced to 12.5 mm, by appropriately altering the aperture setting (i.e., by closing the stop or diaphragm), then the ratio will become 12.5:50 = 1:4, i.e., *f*/4. The aperture with half the diameter has only a quarter of the area of the initial aperture and therefore admits only a quarter of the initial amount of light into the camera (Fig. 2b). The *f*-numbers (aperture settings) of a camera lens are usually so arranged that each next higher *f*-number corresponds to a reduction of about one-half in the light-transmitting power of the lens, e.g.:

aperture (*f*-number)	1.4	2	2.8	4	5.6	8	11	16
relative brightness	1	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{128}$

Hence a lens with *f*/2.8 is sixteen times "faster" than one with *f*/11. If, for instance, an exposure time of $\frac{1}{500}$ sec. is sufficient at *f*/2.8, it will be necessary to give sixteen times as long an exposure at *f*/11, i.e., about $\frac{1}{30}$ sec. For a telephoto lens with a focal length of, say, 100 mm to have the same speed as the above-mentioned normal lens of 50 mm focal length, it will have to have twice the diameter of lens (Fig. 3). This larger lens will transmit four times as much light, but as the image formed on the film is twice as large linearly, i.e., has four times the area of the image formed by the normal lens, the brightness remains the same. With close-up photography (Fig. 4) the distance between the lens and the film will become larger than the focal length of the lens, so that the cone of illumination becomes narrower and the actual light-transmitting effect of the lens diminishes even though the aperture is not altered. If supplementary lenses are used for close-up work, however, the focal length of the lens system as a whole is thereby reduced, but not the distance from lens to film; the actual light-transmitting effect is therefore not reduced. In the case of a compound lens it is not possible to determine the effective aperture merely from the diameters of the component elements and the diameter of the diaphragm. The effective aperture is called the "entrance pupil"; it depends on the diaphragm and forms the boundary of the beam of rays which is concentrated on the film.

Fig. 1

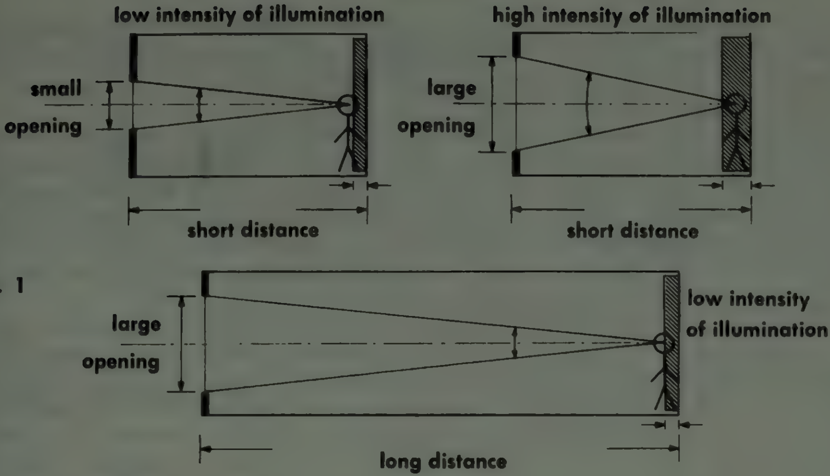
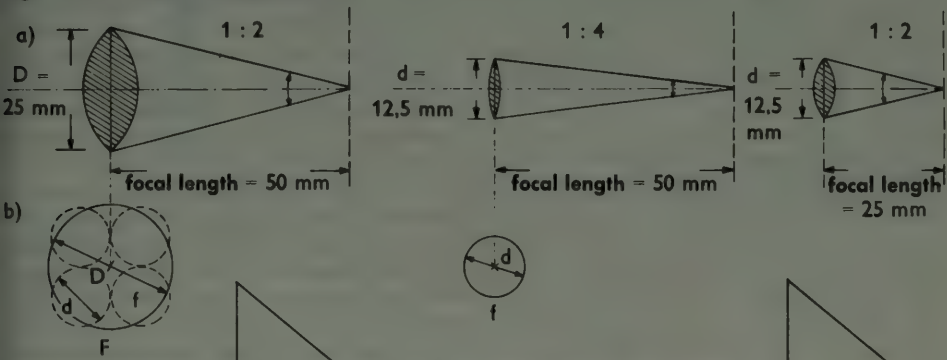


Fig. 2 LENS APERTURE RATIOS



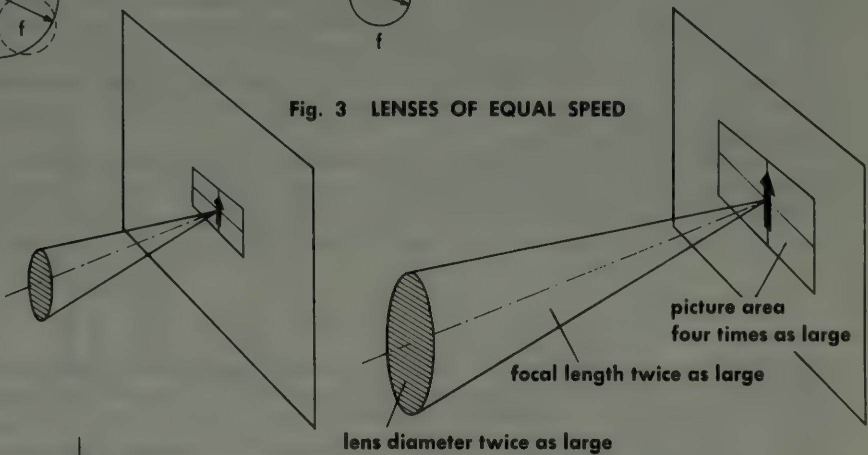


Fig. 3 LENSES OF EQUAL SPEED

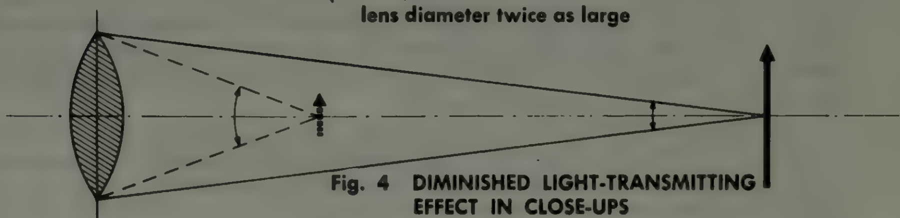


Fig. 4 DIMINISHED LIGHT-TRANSMITTING EFFECT IN CLOSE-UPS

AUTOMATIC EXPOSURE CONTROL

The object of automatic exposure control is to make things easier for the amateur photographer who has no relish for, or interest in, the technicalities of working a camera. He is thus spared the trouble of determining the correct exposure and of selecting and setting the aperture (*f*-number) and shutter speed.

The various systems are all derived from one simple principle: the basis is a built-in exposure meter which detects the brightness of the subject to be photographed. In many 8 mm cine cameras the light received by the meter is transformed into a force which is directly applied to the adjustment of the lightly constructed moving diaphragm. However, in larger standard-size cameras an auxiliary force is needed for adjusting the iris diaphragm and shutter speed. This force is provided either by the user of the camera when he presses the release button, or it is stored up in a spring which is wound beforehand.

When the camera is pointed at the subject, the needle of the measuring device of the exposure meter inside the camera will be deflected a greater or less amount, depending on the brightness. When the release button is actuated, the needle is first arrested by a stirrup bar which presses against it and thereby holds it immovable. This position of the needle determines the exposure value. Then a scanner arm is moved until it touches the arrested needle. At the other end of the scanner are the devices which set the exposure time and the aperture.

The three most frequently encountered methods of automatic exposure control will now be further explained with reference to the somewhat simplified accompanying diagrams:

1. Diaphragm control: To adapt the lens to the brightness conditions, only the aperture is varied. Before the exposure, the lens is set to its maximum aperture. On release, the diaphragm ring is rotated until the toothed stop ratchet engages with a pawl mounted on the scanner. This will occur at an earlier or later instant, depending on the position of the exposure meter needle detected by the scanner arm, and the size of the aperture will accordingly be larger or smaller (alternatively, the scanner may directly control the diaphragm).

The exposure time is always the same. When the user loads his camera with a certain type of film and sets the camera to the film speed (sensitivity of the film), he really selects the shutter speed appropriate to that film. For example, a film of 18 DIN will require an exposure time of $\frac{1}{125}$ sec., while a 15 DIN film, which is only about half as sensitive, will require $\frac{1}{60}$ sec. The exposure time is accordingly adjusted automatically over the aperture range from, say, *f*/2.8 to *f*/22. If a more sensitive film is used, it does not enable photographs to be taken under poorer lighting conditions, but it does enable the exposure time to be reduced (e.g., to $\frac{1}{500}$ sec. for a 24 DIN film), which is essential for photographing fast-moving objects, as in the case of sports events.

2. Programme control of diaphragm and shutter speed: The exposure meter measures the brightness and its needle position is scanned as already described. In this system, however, a so-called programme ring is rotated, which effects continuous adjustment of the shutter speed and the aperture through the agency of control cams and rods; for example:

$$2.8 \quad 2.8 \rightarrow 4 \rightarrow 5.6 \quad 5.6 \rightarrow 8 \quad 8 \rightarrow 11 \quad 11 \rightarrow 16 \rightarrow 22$$

$$\tfrac{1}{30} \rightarrow \tfrac{1}{60} \quad \tfrac{1}{60} \quad \tfrac{1}{60} \rightarrow \tfrac{1}{125} \quad \tfrac{1}{125} \rightarrow \tfrac{1}{250} \quad \tfrac{1}{250} \rightarrow \tfrac{1}{500} \quad \tfrac{1}{500} \quad \tfrac{1}{500}$$

This range of values shows that first the shutter speed is increased (i.e., the exposure time is shortened) as a precaution against blurring of the picture as a result of shaking the camera. Then, when the lens has been stopped down to give better depth of field (see page 192)—at about *f*/8, say—the exposure times are further shortened for photographing fast-moving objects. The film speed has already been taken into account in scanning the exposure meter needle, the meter having first (at the time of loading the camera with that particular film) been appropriately pre-set to that film speed.

3. Exposure time pre-selection and diaphragm control: With this system the user of the camera can choose the exposure time and the aperture for himself. The exposure time ring is coupled to the DIN (or ASA film speed) ring, so that any alteration to the shutter adjustment will automatically also adjust the exposure meter mechanism. The exposure meter needle gives readings on a scale of aperture settings (*f*-numbers) seen in the viewfinder. By altering the exposure time, it is also possible to pre-select the aperture, which constantly adjusts itself to the light conditions.

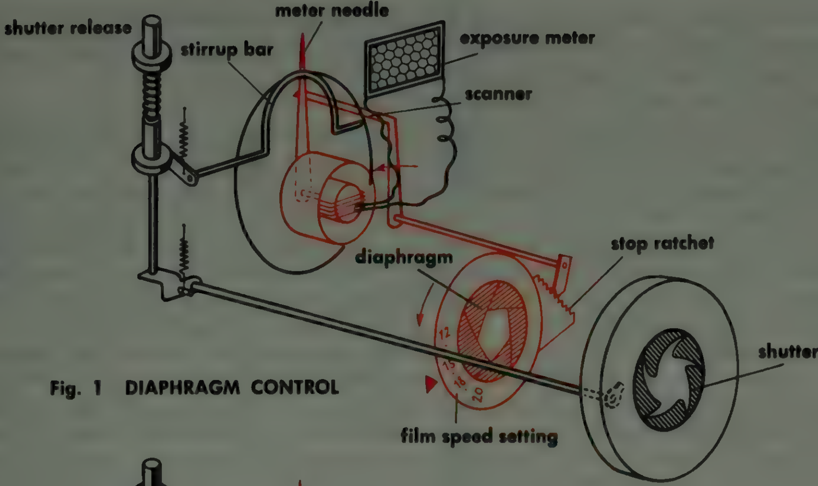


Fig. 1 DIAPHRAGM CONTROL

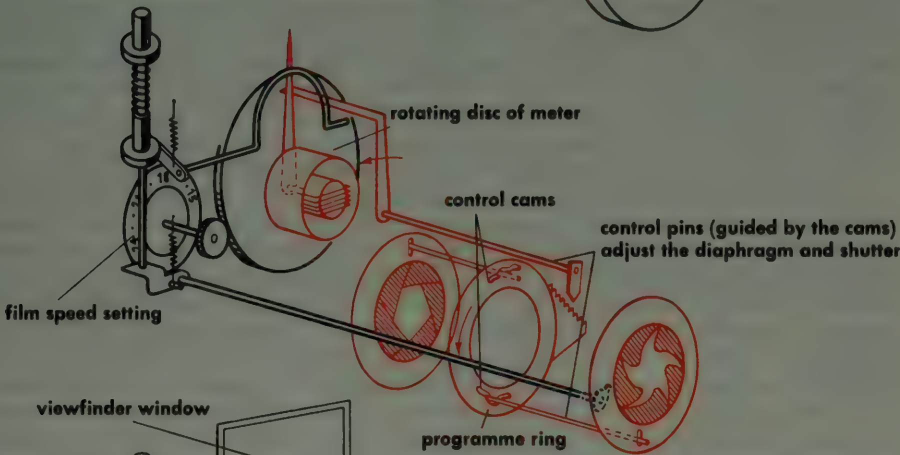


Fig. 2 PROGRAMME CONTROL OF DIAPHRAGM AND SHUTTER SPEED

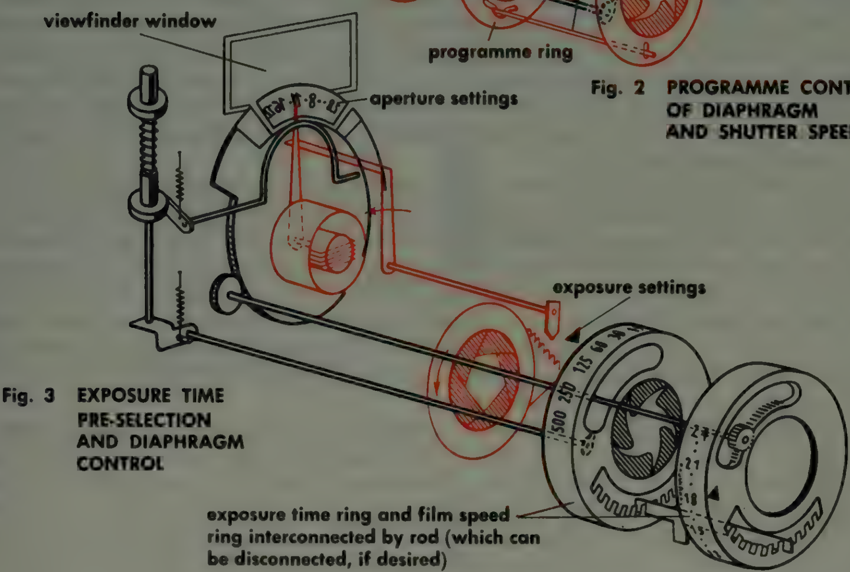


Fig. 3 EXPOSURE TIME PRE-SELECTION AND DIAPHRAGM CONTROL

By depth of field (or depth of focus) in photography is understood the range of distances within which the subject must be located in relation to the camera in order to produce a reasonably sharply focused image.

According to the well-known formula for lenses $\frac{1}{g} + \frac{1}{b} = \frac{1}{f}$ (see page 158), only one object plane (G_e) is sharply focused in the image plane (B_e) (Fig. 1). Objects which are farther away from the lens than the distance G_e have a larger image distance so that the image-forming rays of light converge at a plane B_n behind the image plane B_e to which the lens has been focussed. The image actually formed on B_e (i.e., on the film) will therefore not be in focus but will, instead, be somewhat blurred. Similarly, the image of an object which is closer to the lens than the distance G_e will have a shorter image distance and at a plane B_w in front of the plane B_e. The image actually formed on B_e will therefore likewise be blurred.

In reality, however, there is some latitude. The sharpness of vision of the human eye has its limitations, and some lack of sharpness in the image can be tolerated without being objectionably noticeable as blurring of the picture. Just how much lack of sharpness is acceptable will depend on how much the picture is to be enlarged and on how much importance the observer attaches to distinguishing minute details in the picture. In a sharply focused image each point of the object is reproduced as a point in the image. If the image is out of focus, each point is reproduced as a tiny circle of finite diameter in the image. For amateur photographic purposes a diameter of $\frac{1}{20}$ mm is considered a permissible limit for this so-called "circle of confusion" for 6 cm × 6 cm (pictures and $\frac{1}{30}$ mm for miniature size (24 mm × 36 mm).

The size of the circle of confusion depends on three factors:

1. The distance: The farther a point is located away from the object plane (G_e) on which the camera has been focused, the farther will its image be outside the corresponding image plane (B_e) (Fig. 2). The closer the object* is to the lens, the greater will its image distance vary. For this reason the available depth of field beyond the focused distance is always greater than that in front of the focused distance; for example, when the camera is focused on an object 20 m away, the depth of field may range from 10 m to infinity. With close-ups the depth of field diminishes considerably and may be no more than a few centimetres or indeed a few millimetres.

2. The aperture: When the camera is stopped down (i.e., the diaphragm is closed), the cone of light rays is narrowed and the circle of confusion is reduced in diameter. In Fig. 2 the reproduction of the points III and IV was too blurred, but after the aperture has been reduced (Fig. 3), the corresponding circles of confusion are no larger than those of I′ and II′ previously. (The circles of confusion associated with I′ and II′ have similarly been reduced).

3. The focal length (Figs. 4a and 4b; cf. page 158): As already stated, the image distance varies very considerably when the object is brought close to the lens. An object which is at a distance of, say, 5 m from the camera will be practically located at infinity for a wide-angle lens with a focal length of 35 mm (cf. page 180), whereas for a telephoto lens with a focal length of 500 mm it will be almost in the close-up range. Hence wide-angle lenses and lenses of short focal length have a greater depth of field. Of course, this advantage is cancelled if, in order to obtain the same size of image, the picture subsequently has to be considerably enlarged or the camera has to be brought closer to the subject.

Depth-of-field tables for camera lenses can be based on these considerations. Most lenses have a depth-of-field scale (Fig. 5) on which the apertures (f-numbers) are indicated symmetrically in relation to the focussing mark. For example, if the lens of a miniature camera has been set to $f/11$ and is focussed to 3.3 m, the scale shows that the depth of field extends from 2.2 to 6.8 m, meaning that anything within this range will be within reasonably sharp focus. Conversely, if the subject to be photographed is of some considerable depth, the photographer will rotate the distance-setting ring until the nearest and the farthest point are located opposite equal f-numbers, and he will then set the aperture to this number.

* The term "object" is used in optics generally; more particularly in photography the term "subject" is used to denote the person or thing to be photographed.

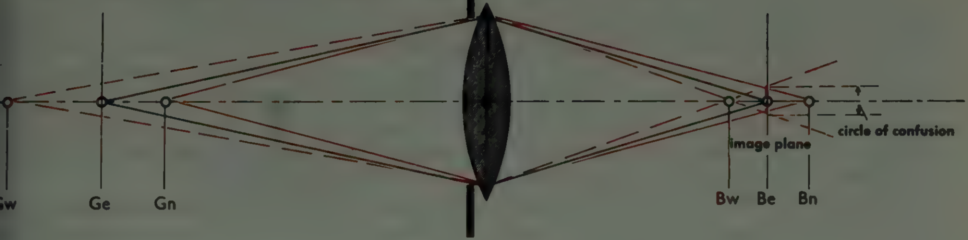
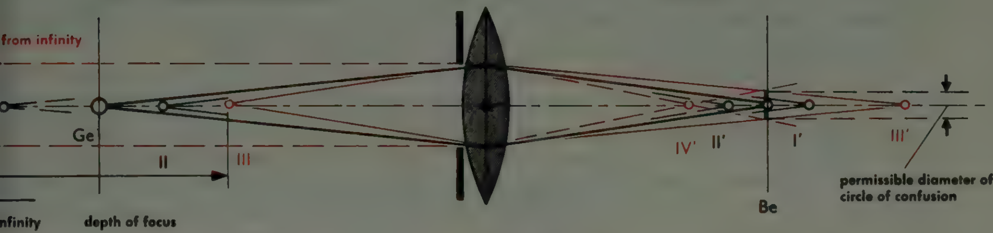


Fig. 1 HOW THE CIRCLE OF CONFUSION IS CAUSED

Fig. 2 DEPTH OF FOCUS AT MAXIMUM APERTURE

Fig. 3 INCREASING THE DEPTH OF FOCUS
BY REDUCING THE APERTURE



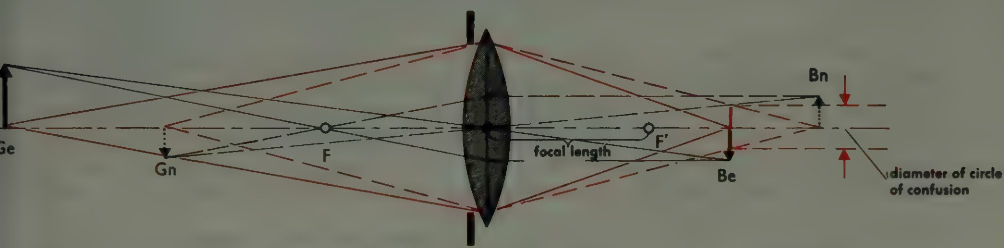
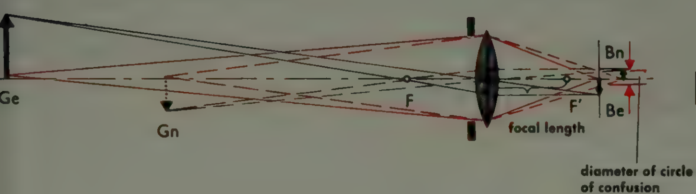


Fig. 4 DEPTH OF FOCUS FOR LONG (a)
AND SHORT (b) FOCAL LENGTH



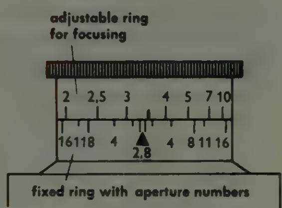


Fig. 5 DEPTH-OF-FOCUS SCALE

In photography, exposure meters are used for determining the correct exposure time for a pre-selected aperture (*f*-number) and a given film speed. Besides optical exposure meters, electric exposure meters are now predominantly used. The light reflected from the subject into the meter is converted into a feeble electric current by means of a photo-electric cell (cf. page 120). This current is measured by the deflection of a needle or pointer, the deflection being larger or smaller according to the brightness of the subject. The precise exposure time (for the given film speed and aperture setting) is indicated on an appropriate scale.

The two principal parts of an electric exposure meter are the photo-electric cell and the measuring device. A device for admitting the light and a contrivance for obtaining the required readings are additional features. The light which enters through a honeycomb lens and a cellular grid strikes the photo-electric cell (the object of the lens and grid is to confine the angle of entry of the light approximately to the angular field of a normal camera lens). The photo-electric cell consists of an iron plate provided with a thin coating of selenium, which in turn is covered with a coating of platinum of such extreme thinness (about $\frac{1}{100000}$ mm) as to be transparent to light. When light strikes the selenium, electrons are released in the latter. These enter the platinum coating and, as a result, an electric current—a very feeble one, it is true—will flow in a circuit connecting the platinum and the iron plate. The light striking the selenium thus produces an electric current; this current is stronger in proportion as the light intensity (i.e., the brightness of the subject) is greater, and vice versa. If a small current measuring instrument is included in the circuit, the pointer of this instrument will provide an indication of the strength of the current. The appropriate exposure times for various aperture settings can then be read from a special scale which is adjustable to the deflection of the pointer. Pre-adjustment to the appropriate film speed (sensitivity of the film) is provided.

Fig. 2 shows the internal parts of a well known type of exposure meter. The adjusting knob is turned until the curved line marked on the roller rotated by this knob coincides with the intersection of the pointer with the reference line. At the same time, bevel gears connected to the adjusting knob cause a flexible band, on which the exposure times and exposure values are marked, to move in relation to the fixed scale of *f*-numbers. The roller with the curve is adjustable to suit the particular speed (sensitivity) of the film used.

In other exposure meters the arrangement of the scales of exposure times and *f*-numbers may be rather different, though the general principle is the same as that described.

Instead of selenium photo-electric cells, so-called photo-resistances are nowadays used in some exposure meters (Fig. 4). A photo-resistance does not convert light into electrical energy but undergoes a change in electrical resistance when light strikes it. The resistance to the passage of current diminishes with increasing light intensity, so that the current—which has to be provided by an auxiliary source such as a small battery—becomes stronger. Meters of this type can be made sufficiently sensitive to respond to very weak light (e.g., moonlight).

The exposure meter is mostly used by pointing it at the subject to be photographed (reflected light measurement). Alternatively, the light that actually illuminates the subject may be measured. This includes the light reflected from adjacent objects. For this reason a so-called incident light attachment or a light diffuser is used, enabling light to be directed into the meter over a much wider angle of entry (Fig. 4).

With regard to built-in exposure meters for automatic exposure control, see page 190.

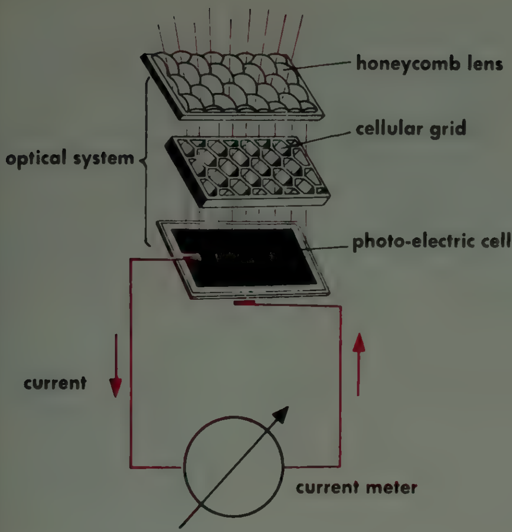


Fig. 1 MAIN FUNCTIONAL PARTS OF AN EXPOSURE METER

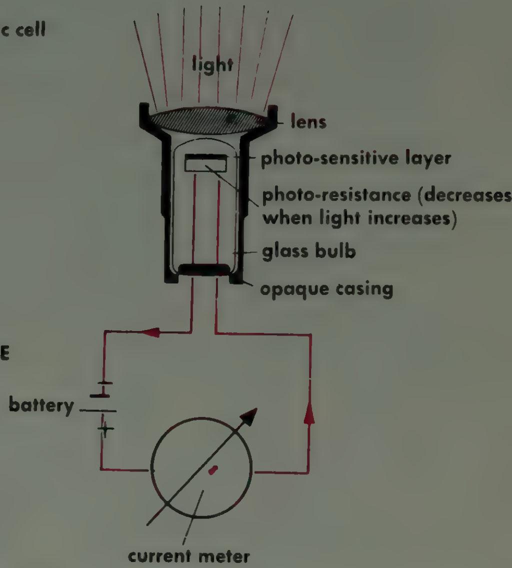


Fig. 3 PHOTO-RESISTANCE (*schematic*)

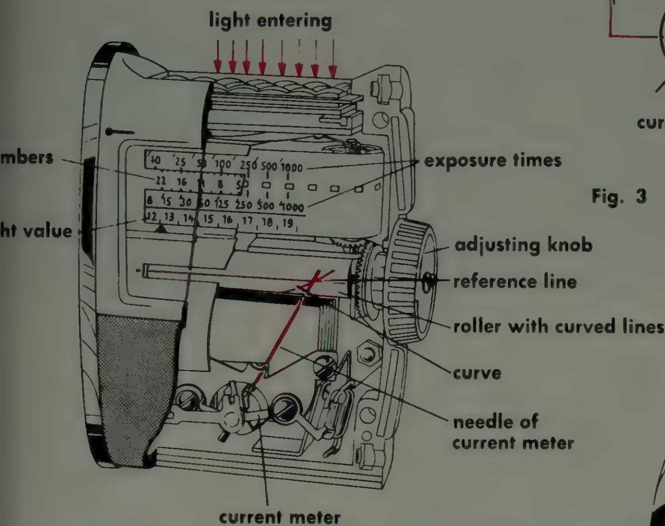


Fig. 2 ELECTRIC EXPOSURE METER (*reflected light measurement*)

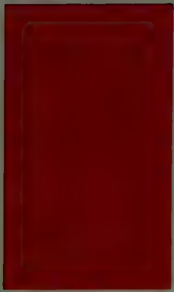
Fig. 4 ELECTRIC EXPOSURE METER (*incident light measurement*)

Colour printing in the present context refers to the production of photo-engraved colour plates by means of the three- and four-colour processes.

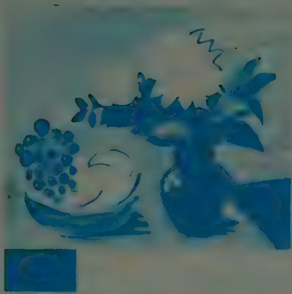
If an area is covered with a mosaic-like pattern of dots in the three primary colours (yellow, red, blue) (Fig. 1c) and is viewed from a sufficiently long distance (so that the individual points cannot be separately distinguished), it appears grey. If numerous dots in two primary colours are placed side by side (Fig. 1c, lower squares) they will, when viewed from a distance, appear as a mixed (or compound) colour. This apparent merging of colours to give the appearance of another colour is one of the fundamental principles of colour printing. The first step in preparing four-colour printing plates from the original painting or photographic transparency is to make colour-separation negatives by means of colour filters on the camera. A different filter is used for each colour negative. Each filter allows only its complementary colour to pass. To make a negative for the blue printing plate, a red filter is used; for the yellow printing plate, a blue-violet filter; for the red printing plate, a green filter. A pale yellow filter is used for obtaining the black values. These various colour negatives are shown in Figs. 1b, 2b and 3b; the colours of the corresponding filters are shown in Figs. 1a, 2a and 3a. The purpose of the black plate (Fig. 4) is to cover certain white areas of the base (i.e., the paper on which the picture is printed) and to give depth and detail to the picture.

The negatives are developed, printed on metal, and etched in much the same way as when making a monochrome halftone plate. The plates are inked and printed in the sequence shown (Fig. 1–Fig. 4), each being superimposed exactly upon the other. The picture is successively thus built up step by step, as shown in Figs. 2c, 3c and 5. The surface of each printing plate is composed of a large number of dots, these being formed by a special process. When the four colours from the plates are superimposed, the resulting print contains not only the four colours, but also the compound colours formed by the blending of these. Fig. 6 is an enlarged detail of Fig. 5 and shows how this effect is obtained.

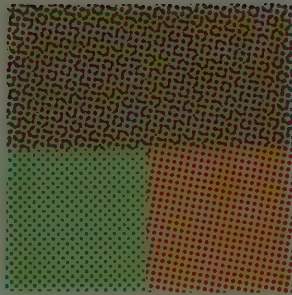
To achieve good results, it is of course essential to make the successive coloured impressions register exactly one upon the other. The coloured inks must be clear and clean, and variations in the shade and intensity of the colours must be avoided. In some cases, to obtain very high quality, it may be necessary to employ auxiliary colours. Sometimes, too, the colour negatives may have to be retouched by hand.



1a



1b



1c

2a

2b

2c

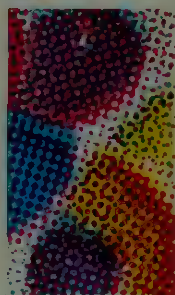
3a

3b

3c

4

5



6

A certain type of colour television camera contains three image orthicon tubes (see page 146). By means of a system of mirrors and colour filters the first tube forms a red image (R), the second forms a green image (G), and the third forms a blue image (B). The three camera tubes have essentially identical scanning patterns, so that the picture signals developed by the respective tubes represent images which are identical except that they differ in colour. (See Fig. 1).

By means of an electric transmission system the primary colour signals E_R, E_G and E_B are simultaneously fed to three colour picture tubes and converted back into three separate colour images (red, green and blue). By means of a system of colour-selective (dichroic) mirrors the viewer sees the three pictures as one superimposed picture in which the three colours are blended to give additively mixed colours, just as in colour printing (see page 197).

In the system developed by the National Television System Committee (NTSC system) in the United States, the E_R, E_G, E_B signals (primary colour signals) are converted by a device called a colour coder into a luminance (i.e., brightness or "brilliance") signal, E_Y and a chrominance signal. Chrominance comprises two independent characteristic quantities: "hue" and "saturation". The luminance signal can be received by an ordinary (monochrome) television receiver and produces a black-and-white picture. This technique comprising a luminance and a chrominance transmission is known as compatible colour television. The luminance signal is subtracted from the primary colour signals, and the colour-difference signals thus obtained are then further combined to produce two signals E_I and E_Q, which are then mixed with the chrominance subcarrier signal. This signal is amplitude-modulated in accordance with the saturation values and phase-modulated in accordance with the hues. The luminance and chrominance components are combined to form the overall colour picture signal, which is then transmitted. The picture signal wave is a composite wave in which the chrominance wave is superimposed upon part of the luminance wave.

(Continued)

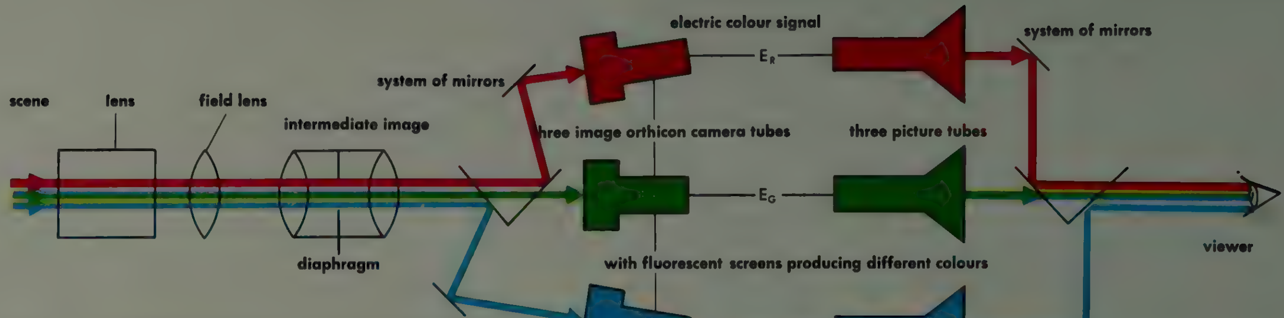
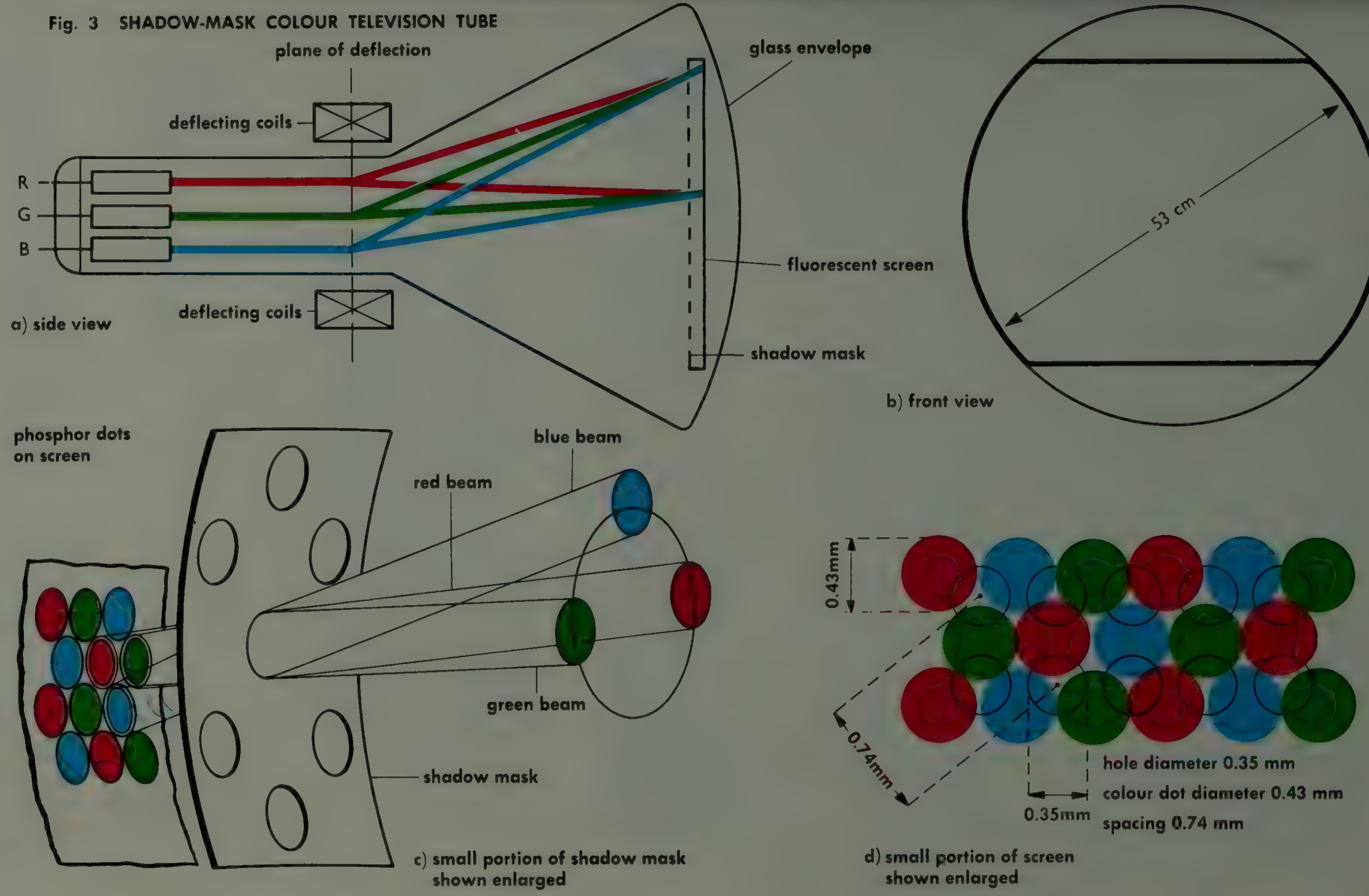


Fig. 1 SIMULTANEOUS SYSTEM OF COLOUR TELEVISION

The primary colour signals, which have been recast into luminance and chrominance components at the transmitter have to be reconverted into primary colour signals at the receiver before they can be applied to the colour picture tube. Instead of using three picture tubes as in the system outlined above, the NTSC system uses only one picture tube, known as a shadow-mask tube, which contains three electron guns which produce three separate electron beams, which move simultaneously in the scanning pattern over the viewing screen and which respectively produce a red, a green and a blue image. The screen is composed of three separate sets of uniformly distributed phosphor dots. The dots of each set glow in a different colour. Electrons discharged by the gun controlled by the red primary colour signal impinge only on the red-glowing phosphor dots and are prevented from impinging on the green- and blue-glowing dots by a mask which contains about 200,000 tiny holes, each of which is accurately aligned with the different coloured phosphor dots on the screen. Similarly, the electrons from the two other guns fall only on the green and the blue dots respectively. In this way three separate primary colour images are formed simultaneously. The dots producing the three different colours are so small and so close together that the eye does not see them as separate points of light.

In the colour television receiver the luminance component is applied simultaneously to all three electron guns of the picture tube. The receiver contains circuits which perform the inverse operations of the addition and subtraction circuits at the transmitter. In this way three colour-difference signals are obtained (the difference between the luminance signal and the primary colour signals). The three colour-difference signals are applied to the respective electron guns in addition to the luminance signal. The net control signal applied to each gun corresponds to the primary colour signal coming from the respective camera tube.

Fig. 3 SHADOW-MASK COLOUR TELEVISION TUBE



COLOUR PHOTOGRAPHY

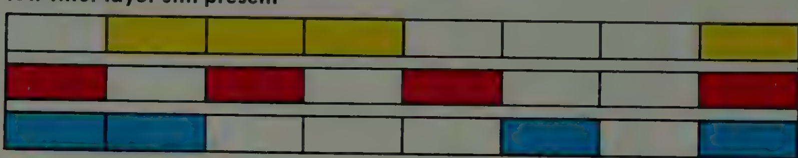
There are many different processes of colour photography, which differ in detail but they can be divided into two main groups: additive processes (for transparencies only) and subtractive processes (for transparencies and also for paper prints). In all cases the fundamental principle is the same: images in three primary colours must be obtained which, in combination with one another, produce the desired colour picture. The individual processes differ in that, for example, the subject may be photographed three times in succession in three different colours (e.g., yellow, green, red) or that, alternatively, the three films are exposed simultaneously in a special camera fitted with a system of lenses and mirrors for achieving this result, or again that only one film is used, which is provided with three superimposed layers of emulsion, each sensitive to a different colour. The principles whereby the three single-colour images are combined to produce the final result also varies from one process to another. The reversal process and the negative-positive process described here belong to the group of subtractive processes, which are now used almost exclusively for colour prints on paper and transparencies for still and motion-picture projection.

The emulsion layers incorporate chemicals which, on exposure to light, form dyes to a degree corresponding to the amount of light. Most modern colour photographic processes embody this principle of dye formation, which is based on the fact that when a developer reacts with silver bromide to reduce it to silver, the resulting oxidised developer can react with certain dye-forming chemicals (called couplers) to form dyes. The developers used are diamines.

In a well-known type of colour film the film is coated with three emulsion layers which respond respectively to blue, green and red light. A yellow filter layer is inter-

a) Colours in the subject to be photographed

b) Layers in the film after the first (black and white) development: black grains represent developed silver, light grains represent residual silver bromide, yellow filter layer still present



c) The same film after processing: colour developed at the residual silver bromide, silver and yellow filter layer dissolved out

d) Colours in the finished film

posed between the blue-sensitive top layer and the two other layers in order to prevent blue light from reaching these. Positive pictures are obtained directly by reversal. First, the film is developed to give a negative in all three layers, followed by exposure to red light through the base, and again developing; then exposure to blue light through the front, again followed by developing. The silver in all three layers is bleached out, leaving the dyes to form the colour picture (see illustration below).

In another process, called the negative-positive process, the film is coated with red-, green- and blue-sensitive layers. After exposure, the film is processed with a developer which produces silver plus dye images in each layer. The silver is removed, and a colour negative is obtained. The dyes formed in each layer have the colour complementary to the colour of the light that caused it to be formed. Thus, blue light produces a yellow dye, green light a purple dye, and red light a blue-green dye. Thus all the colours in the negative are complementary to those in the subject photographed. The negative is printed on to paper coated with a similar set of emulsion layers, so that prints are obtained in which the original colours of the subject are reproduced (see illustration). In a refinement of this process the film has an extra layer in which a black-and-white negative can be developed. This acts as a mask and improves the colour quality of the print.

The coloured diagrams show how the final colour reproduction is obtained. If small blue and yellow dots are situated close beside one another, the overall visual impression is of a green colour. This is the principle of additive colour blending. However, the final transparencies or prints obtained in colour photography comprise superimposed single-colour layers which are partly transparent. The layers act as filters, so that only those rays of light which are not absorbed by these layers reach the eye. This is the principle of subtractive colour blending.

a) Colours in the subject to be photographed

b) Layers in the finished negative

c) Finished negative: colours are complementary to those in the subject

d) Layers in the finished positive

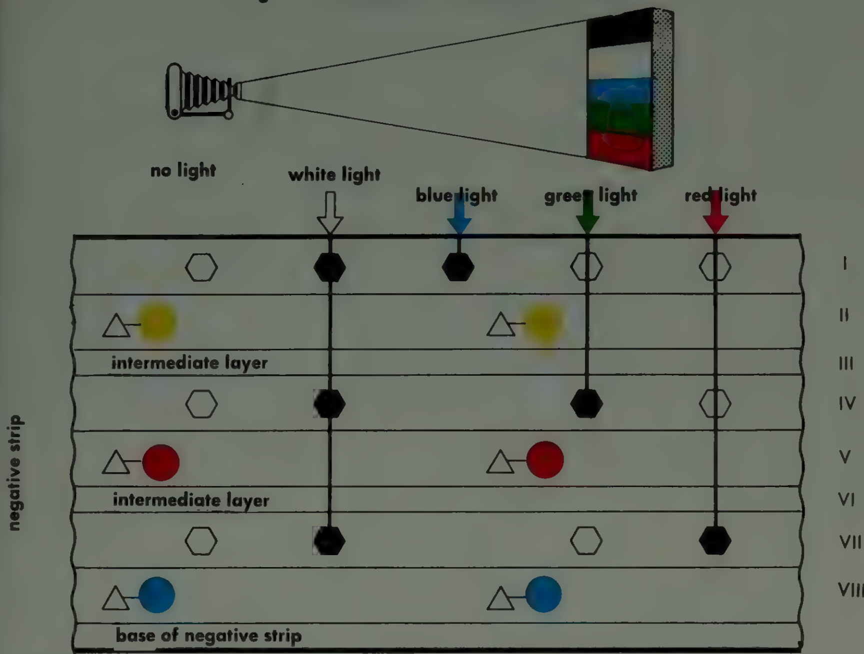
e) Colours in the finished positive

The processing of ordinary colour films to produce the final result—the colour positive—is a fairly laborious procedure, involving a number of operations performed under carefully controlled chemical and temperature conditions and with meticulous timing.

On the other hand, the colour photography process developed by Dr. Land produces a colour positive in only one minute. The film is coated with three colour-sensitive layers containing silver halide crystals, three layers containing dyes and developers, and two intermediate layers. Blue light acts upon the blue-sensitive layer I; green light passes unhindered to the green-sensitive layer IV; red light passes through both these layers to the red-sensitive layer VII. Now a strip is pulled out of the camera, causing a viscous developer-activating paste to be squeezed out of a capsule, while steel rollers press the negative layer (with the squeezed-on paste) firmly on to the so-called positive layer. The latter comprises the colour reception layer, an intermediate layer, and an acid layer for neutralising and stabilising the colour image. The activating paste diffuses quickly into the negative layer and first enters the blue sensitive layer I, where it develops the silver halide crystals; then it liberates a yellow dye in layer II, passes through the intermediate layer III, develops the silver halide crystals in the green-sensitive layer IV, liberates a purple dye in layer V, passes through the intermediate layer VI, develops the silver halide crystals in the red-sensitive layer VII, and finally reaches the layer VIII, where a blue-green dye is liberated. All the dye molecules liberated by the paste from the compounds in which they were previously held in a latent form are now free to diffuse in all directions. Where they encounter a grain of silver (into which a silver halide particle which was exposed to light has been converted by the developing action) they are retained and held by it. Those dye molecules which encounter no silver grains continue their migration until they reach the positive layer, where they, too, are held.

(Continued)

Fig. 1 EXPOSURE OF EIGHT-LAYER FILM



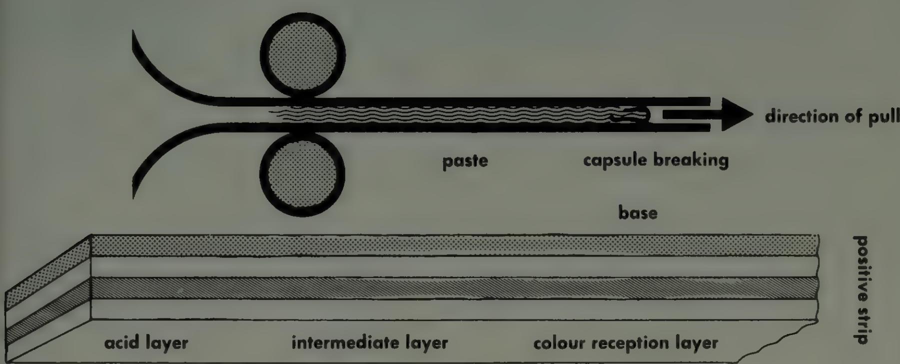


Fig. 2 THREE-LAYER POSITIVE STRIP IS PRESSED AGAINST NEGATIVE STRIP; DEVELOPER-ACTIVATING PASTE SERVES AS "GLUE"

How is the colour image produced? Where red light has fallen on the film, the blue-green dye from layer VIII is held by the silver grains in layer VII. Since the red light had no effect on the layers IV and I (these being sensitive to green and blue respectively) the purple and the yellow dye respectively liberated in the layers V and II can pass freely to the colour-holding positive layer. Yellow and purple mixed produce the red colour in the positive layer. Blue light affects only the blue-sensitive layer I, and the yellow dye formed in the adjacent layer is retained and held by the grains of silver in layer I. But now the purple dye from layer V and the blue-green dye from layer VIII can pass unhindered to the positive layer, where the two dyes are held: purple and blue-green mixed produce blue. Similarly, green light affects on the green-sensitive layer IV; the silver grains in that layer retain and hold only the purple dye, whereas the blue-green and the yellow dye travel unhindered to the positive layer, where their mixture produces green. In the case of white light all the dyes are retained, because this light affects all three colour-sensitive layers. In the shadows ("black"), no dye is retained, and all three dyes reach the positive layer: their mixture produces black. Finally, water is formed in an alkali-neutralising process and serves to wash the positive. The process is illustrated schematically in the accompanying diagrams.

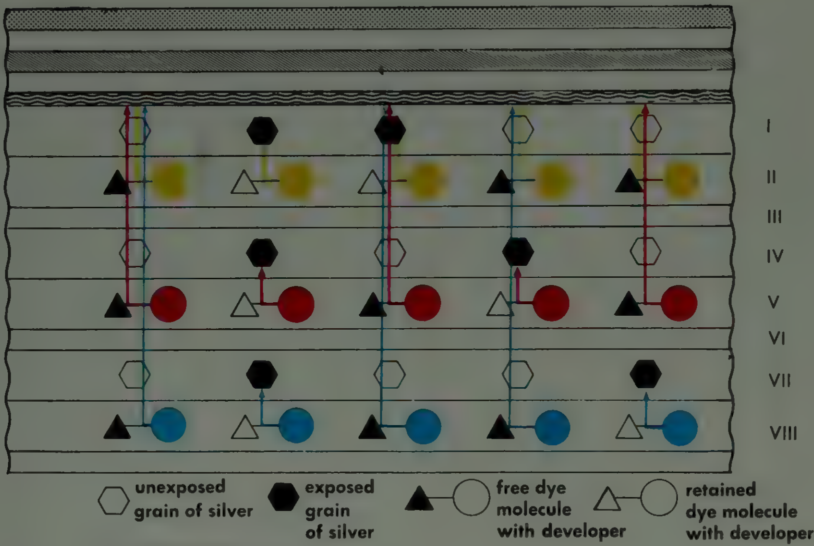


Fig. 3 DEVELOPMENT PROCESS

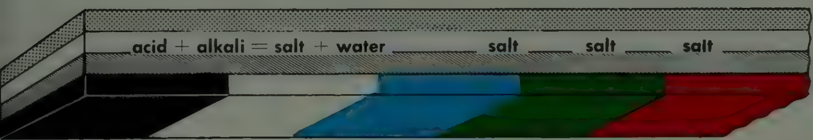


Fig. 4 FINALLY DEVELOPED POSITIVE

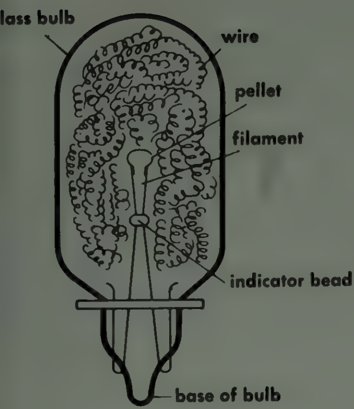
Flash bulbs are required to produce the greatest possible light flux (measured in lumens) in the fraction of a second during which the shutter of the camera is open. The total quantity of light emitted by a bulb during this period is measured in lumen-seconds. The light is produced by the combustion of a coil of aluminium-magnesium wire or zirconium wire with oxygen. An electric current impulse from the flash contact in the camera causes a tungsten filament to glow and ignite small pellets of a substance similar in composition to the head of a match, which in turn ignite the metal wire coil. The oxygen filling in the bulb has a pressure of less than 1 atm., which ensures that the glass bulb will not shatter as a result of the increase in pressure due to the explosion-like combustion process. As a further precaution the bulb is coated internally and externally with a transparent lacquer which prevents glass splinters from flying about. For colour photography there are flashbulbs with a blue-tinted lacquer which ensures that the light emitted has a spectrum composition more or less similar to that of daylight. Inside the bulb is an indicator bead of a blue cobalt salt which turns pink as the result of penetration of atmospheric moisture into the bulb if the latter should become defective by developing a leak either during manufacture or during subsequent handling.

The light-time curves (Fig. 2) give information on the light efficiency and the application purpose of flashbulbs. Of especial importance to synchronisation with the camera shutter is the peak value., i.e., the length of time between the closing of the contact and the maximum light intensity. For ordinary type M flashbulbs it is 18 milliseconds. Most diaphragm shutters have "X synchronisation": contact is made as soon as the shutter is fully open (Fig. 3a). In order that the entire "firing" process of the flashbulb is completed while the camera lens is open, the shutter speed must be at least 30 milliseconds (about $\frac{1}{30}$ sec.). For use with faster shutter speeds many cameras are additionally equipped with a so-called "M" contact, which comes into operation 16 milliseconds before the lens is fully open, so that the shutter opens only a short time before the peak value is reached and only that part of the firing process which corresponds to maximum flash intensity is utilised (Fig. 3b). For comparison, the light-time curve for an electronic flash (see page 210) is also shown in the two diagrams: with "X" synchronisation it is possible to use diaphragm shutters up to $\frac{1}{500}$ sec.; in that case, if the contact is inadvertently set to "M", no light from the flashbulb will reach the film.

With focal-plane shutters the shutter action takes about $\frac{1}{20}$ to $\frac{1}{60}$ sec. to accomplish and shorter exposures can be obtained only by narrowing the slit (see page 186). For this reason the flashbulb employed should have a wider firing curve, so as to ensure that the light intensity on the film at the moving slit is constant over the entire width of the picture.

If the exposure time is longer than the firing time of the flashbulb, the exposure will depend solely on the aperture setting. The aperture (f-number) is calculated from the "guide number" for the type of film used, the formula being: f-number = guide number divided by distance (in metres). To understand the significance of the guide number it must be borne in mind that the intensity of illumination decreases in proportion to the square of the distance from the light source and that the speed (transmitting power) of a lens likewise decreases with the square of the f-number. The two quantities are linked by the guide number whose square is proportional to the quantity of light. For instance, the intensity of illumination at a distance of 5 m is only a quarter of that at 2.5 m. The transmitting power of the lens at $f/4$ is four times as high as that at $f/8$. This means that $f/4$ at 5 m gives the same exposure as $f/8$ at 2.5 m. Both settings can be used with a flash having a guide number of 20. Other possible settings in this case are $f/2.8$ at 7 m or $f/5.6$ at 3.5 m.

Ignition used to be effected by directly short-circuiting the battery through the flashbulb (Fig. 4a). Nowadays small dry batteries, usually of 22.5 volts, are used which charge a condenser when the flashbulb is inserted (Fig. 4b); because of the high resistance in the charging circuit, this current remains so feeble that it does not ignite the bulb. When the shutter is released, the contact in the camera connects the condenser directly to the bulb, and the sudden discharge causes ignition of the latter.



g. 1 FLASHBULB

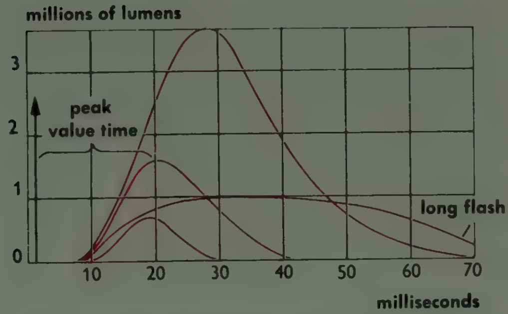


Fig. 2 LIGHT-TIME CURVES OF FLASHBULBS

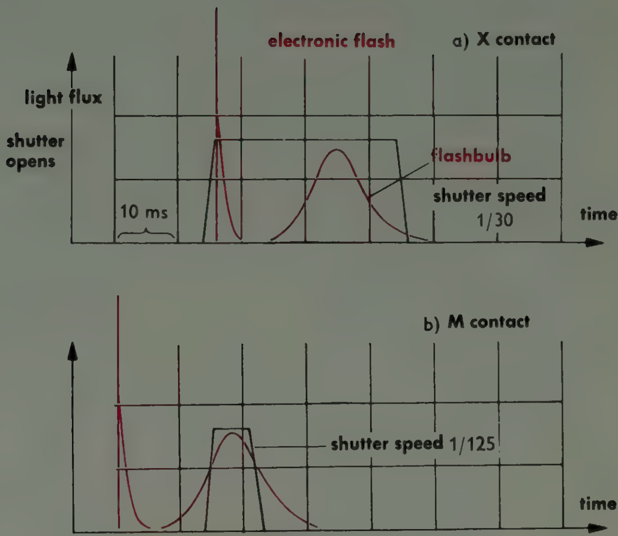
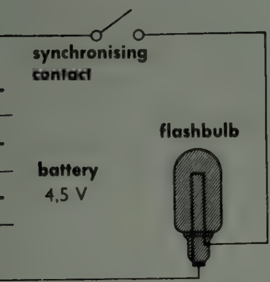


Fig. 3 FLASH SYNCRONISATION WITH DIAPHRAGM SHUTTER



ig. 4a CIRCUIT DIAGRAM FOR IGNITION OF FLASHBULBS DIRECT FROM THE BATTERY

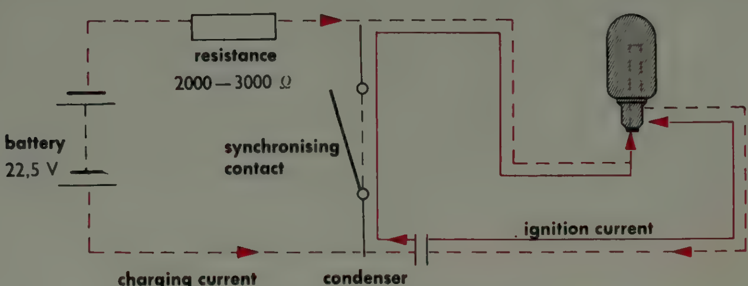


Fig. 4b CIRCUIT DIAGRAM FOR IGNITION OF FLASHBULBS THROUGH A CONDENSER

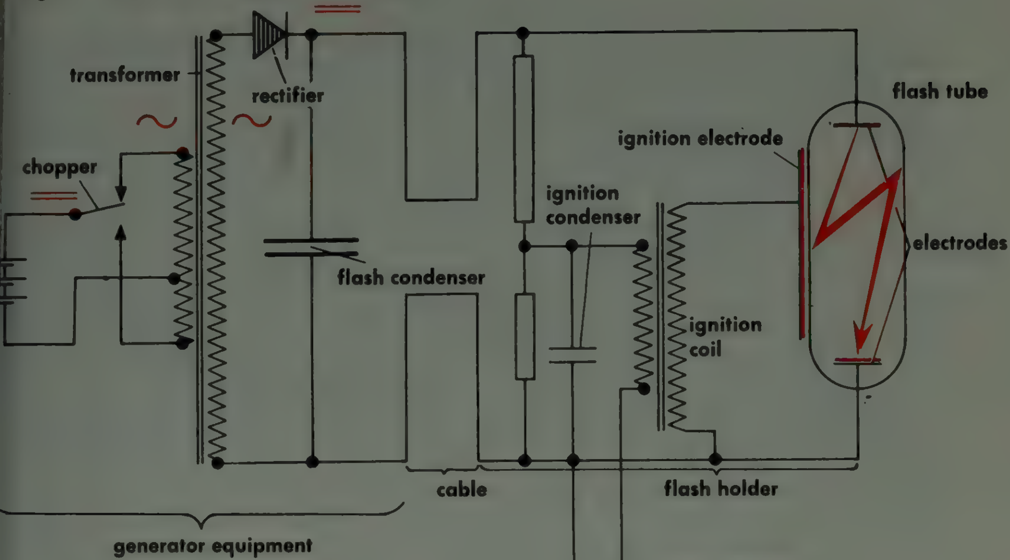
The "electronic flash" used in modern flashlight photography produces a high-intensity source of light of very short duration which has much in common with a flash of lighting. In a thunderstorm, lightning is caused when a cloud becomes electrically charged, for example, in relation to the earth; the cloud and the earth can be conceived as the plates of a giant condenser. If the voltage becomes too high, a flash-over occurs: the gas of the atmosphere becomes temporarily conductive to electricity, and a tremendous surge of current occurs which lasts a very short time. In the electronic flash this effect is simulated by discharging a condenser through a gaseous atmosphere. The apparatus must therefore generate the requisite high voltage and enable the flash to be produced at the desired instant (Fig. 1). Older types of apparatus operated at some thousands of volts, but the more recent ones usually operate at 500 volts. Portable sources of electric current (dry batteries or accumulators), however, supply direct current at very low voltages. Only an alternating current can be stepped up to a higher voltage by means of a transformer. The direct must therefore first be converted into an alternating current. This is done either by means of a so-called "chopper" which changes the polarity (i.e., the direction of flow) of the current about 200 times per second by mechanical action or by means of a transistor circuit which achieves a similar effect (in modern equipment the latter method is almost exclusively employed). The alternating current thus obtained is raised to the required high voltage by means of a transformer (see page 110). A rectifier then converts it back into direct current.

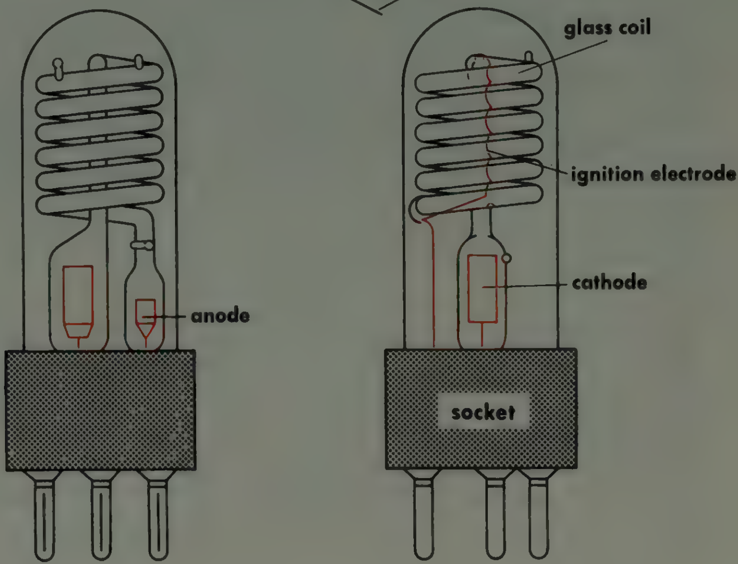
This high-voltage direct current charges a condenser whose two poles are connected to the electrodes of the flash tube. The flash tube and ignition device are in the flash holder. Also, there are modern very small electronic flash sets in which the generator equipment and the reflector are combined into a single unit. The flash tube is filled with an inert gas, usually xenon, which does not conduct electricity. The final voltage at the condenser, and therefore at the electrodes, is attained after a charging time of about 3 to 15 sec. This voltage is not so high, however, that it can by itself initiate a discharge through the flash tube. When the camera shutter is released, the synchronising contact of the shutter causes a small ignition condenser to discharge through one winding of the ignition coil. As a result of this, a high voltage (about 10,000 volts) is induced at that instant in the other winding. This high voltage impulse, which is applied between the ignition electrode and the negative electrode of the flash tube, ionises the inert gas and makes it conductive to electricity, so that the condenser can discharge itself through the tube. During a period of about $\frac{1}{1000}$ sec. the stored-up electrical energy flows through the flash tube as a high-intensity current (of about 100 amps and more), during which time it causes the gas to glow brightly.

The light emitted by the electronic flash has approximately the same spectrum composition as daylight. For colour photography the usual daylight colour film should accordingly be used.

For synchronisation with the shutter of the camera and for calculating the aperture setting from the "guide number", see page 208.

Fig. 1 ELECTRONIC FLASH (*schematic*)





HIGH-VOLTAGE FLASH TUBE

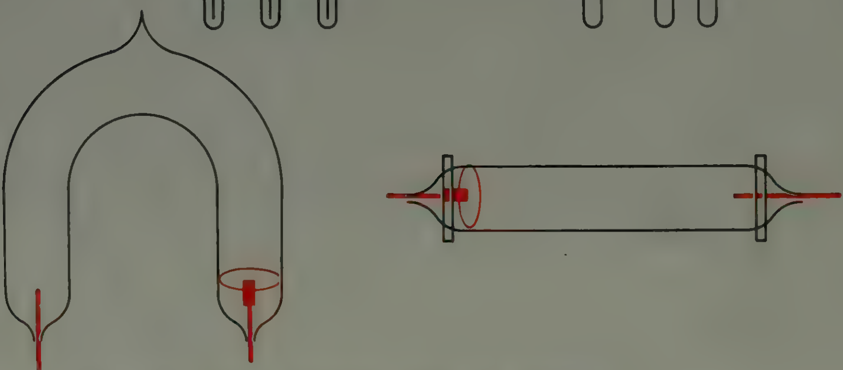


Fig. 3 FLASH TUBES USED IN MODERN AMATEUR EQUIPMENT (*drawn to a much larger scale than Fig. 2*)

When subjects close to the camera are photographed with large apertures (low f-numbers), the depth of field (see page 192) is very small. A mere rough estimate of the distance is generally too inaccurate, and for this reason high-class cameras are equipped with built-in rangefinders. The principle of the rangefinder is rather like that of binocular vision: when an observer looks at a very distant object, his two eyes look in parallel directions; if the object comes nearer, the eyes turn inwards, and they do this more and more in proportion as the distance to the object decreases. The angle between the lines of sight is therefore a measure for the distance (Fig. 1).

Similarly, the rangefinder has two "eyes" or sight openings: one of these is the normal viewfinder, whose optical axis is accurately parallel to that of the camera lens; the other is located at the other end of the "base" and produces a second image (the rangefinder image). The rays of light which form the rangefinder image are reflected by a mirror to the viewfinder, where it is superimposed upon the viewfinder image by means of a semi-transparent mirror (Fig. 2a). In the eyepiece of the viewfinder the photographer thus sees two images, one over the other. This is known as coincident image (Fig. 2b). If the two rays entering the viewfinder and the second "eye" of the rangefinder are parallel (i.e., if they both come from a very distant object), the two superimposed images will accurately coincide and thus appear as one image. An object close to the camera will, however, produce two images which are displaced in relation to each other. To bring them into exact coincidence, the second ray will have to be deflected through a certain angle until it, too, passes through the object. The magnitude of this angle is a measure of the distance to the object in question. The angular rotation of the component which produces this deflection can therefore be coupled to a distance scale (in metres or feet) or directly to the focusing mechanism of the lens. The deflection may be produced by means of a swivelling mirror (Fig. 2a). Greater precision is attainable with the swivelling wedge type rangefinder (Fig. 3) because here the deflecting component has to be rotated through a substantially larger angle to produce a given deflection. In this device two cylindrical lenses, one of which rotates within the other, form a prism of variable refractive power. In Fig. 3 the base with the two fixed mirrors consists of a glass rod for greater strength. In other forms of construction the pair of cylindrical lenses (or sometimes only one swivelling cylindrical lens) is mounted between the mirrors. In the rotating wedge rangefinder (Fig. 4) two circular wedges can be rotated 180° in relation to each other in both directions. In one extreme position they together form a double wedge; in the intermediate position they form a plano parallel plate; and in the other extreme position they form a double wedge which deflects light in the opposite direction. Because of its complicated mechanism, this type has now largely been superseded by the simpler lever-action rangefinders. For the same reason the so-called split-image rangefinder (Fig. 5a) has now largely disappeared, although it is a very convenient one to use because of the clear image it produces in the rangefinder. In this type the image is split into a top and bottom half which are displaced in relation to each other and which have to be brought into coincidence with each other (Fig. 5b).

Nearly all mirror-reflex cameras are now equipped with a focusing aid which, because of its similar action, is often referred to as a "split-image rangefinder". This device comprises two clear glass wedges which are so cemented in position that their point of intersection is accurately located in the plane of the focusing screen. If the image has been sharply focused on the screen, the wedges have practically no effect; the farther the image is away from this plane, the more the two parts of the image are displaced in relation to each other as a result of the deflecting action of the wedges.

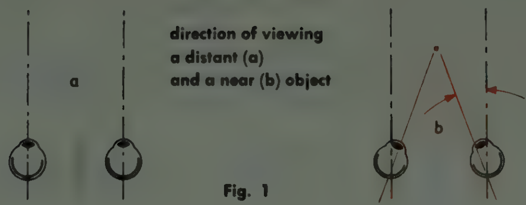
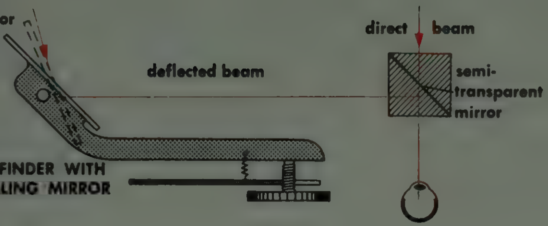


Fig. 1



. 2a RANGEFINDER WITH SWIVELLING MIRROR

Fig. 2b

coincident image

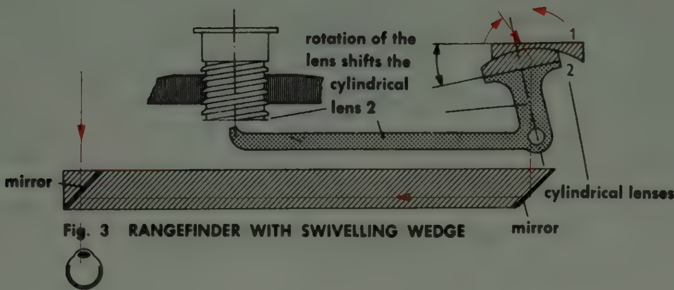


Fig. 3 RANGEFINDER WITH SWIVELLING WEDGE

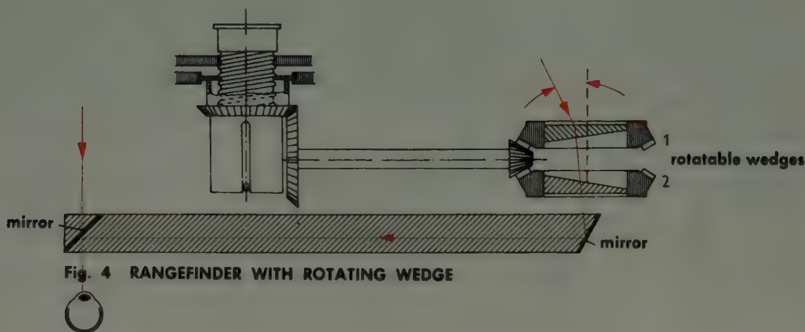
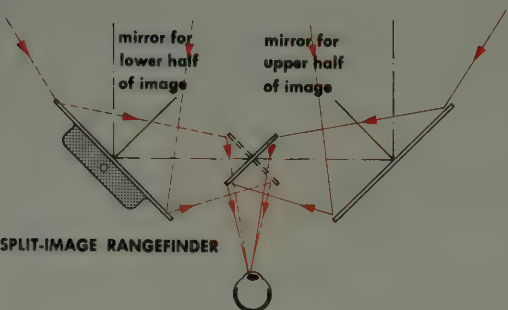


Fig. 4 RANGEFINDER WITH ROTATING WEDGE



g. 5a SPLIT-IMAGE RANGEFINDER

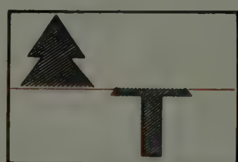


Fig. 5b

split image

There is no difference in principle between an ordinary "still" camera and a cine camera, except that the latter takes a large number of pictures in rapid succession, so that when these are appropriately projected on the screen, an impression of continuous motion is produced.

Fig. 1 shows how the film is run through the camera. The feed spool carries the unexposed film, which is unwound from this spool by the sprocket wheel and moved along to the film gate. The sprocket is driven by a small electric motor or a spring motor. Guide plates keep the film engaged with the sprocket. A pressure plate keeps the film perfectly flat as it moves past the film gate. After exposure, the film is advanced a distance corresponding to one image by the feeding claw. As this movement occurs intermittently, whereas the sprocket rotates continuously, the film must comprise relieving loops. While the feeding claw pulls the film along, the rotary disc shutter covers the film gate and then exposes it for a certain length of time (about $\frac{1}{32}$ to $\frac{1}{50}$ sec.). The feeding claw and the shutter must therefore be interadjusted.

The feeding claw is secured at its pivot; its lever is movable (Fig. 2). The hook engages with the perforation of the film and moves the latter forward a certain distance, as the other end of the claw performs an eccentric motion. In narrow-gauge cine cameras the film is held stationary for as long a time as is necessary for the exposure. The rotary disc shutter therefore performs the function of covering the film—i.e., preventing light entering the camera—during the film advance movement. It consists of a circular disc (Fig. 3) with cut-out sectors. With 16 exposures per second a sector of 180° corresponds to an exposure speed of $\frac{1}{32}$ sec. Fig. 4 shows the interrelated operations of covering the film gate and the film advance movement.

The various running speeds of the camera are adjusted by means of a small governor (Fig. 5). Small centrifugal weights attached to a plate are slidably mounted on a shaft. In the neutral position (Fig. 5a) a spring pushes the plate to the right. Now if the shaft rotates (Fig. 5b) the weights are flung outwards by centrifugal action and pull the plate to the left until it strikes a stop. The position of this stop therefore determines the speed of rotation of the shaft, which will be higher according as the stop is farther to the left.

Slow-motion and time-lapse effects are determined by the ratio of the speed of shooting the film to the speed of projecting it. If the film is run through the camera at a lower speed (i.e., a lower footage per second) than through the projector, the movements of the subject which has been filmed will be seen speeded up. Conversely, slow-motion is obtained by running the film through the camera at a higher speed than through the projector.

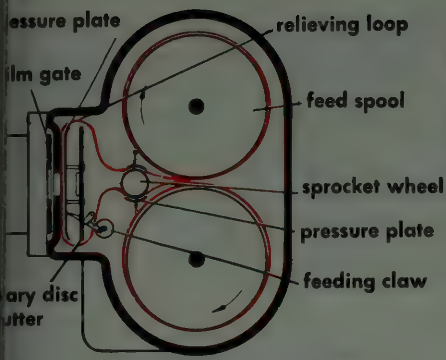


Fig. 1 SCHEMATIC DIAGRAM OF A CINE CAMERA

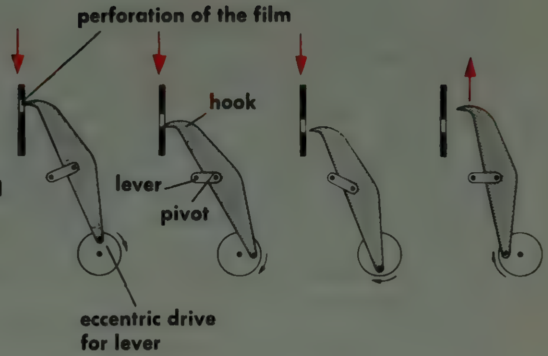


Fig. 2 HOW THE FEEDING CLAW WORKS

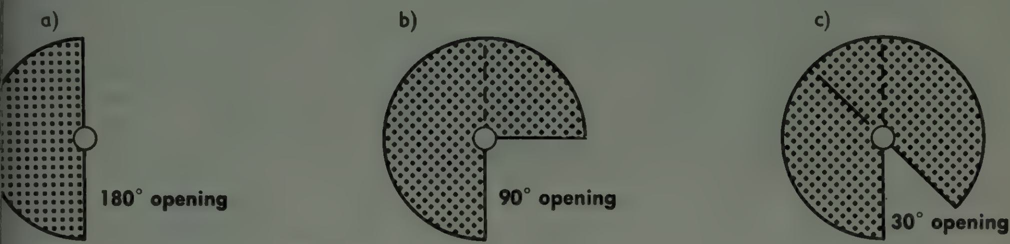


Fig. 3 ROTARY DISC SHUTTER

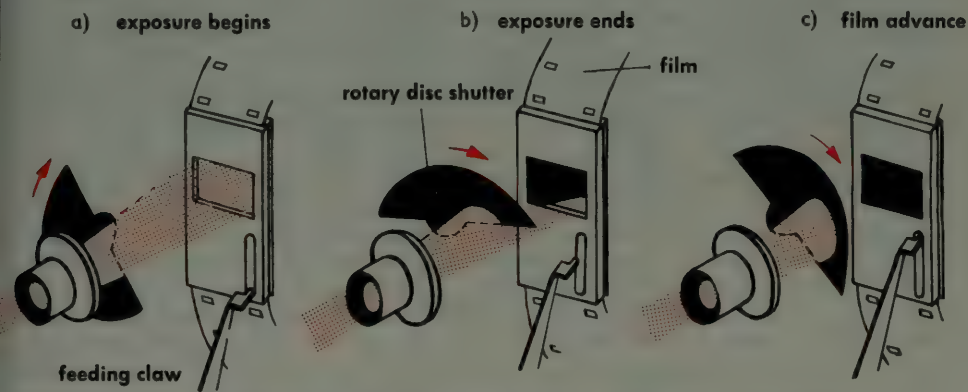


Fig. 4 CO-ORDINATION OF SHUTTER OPERATION AND FILM ADVANCE

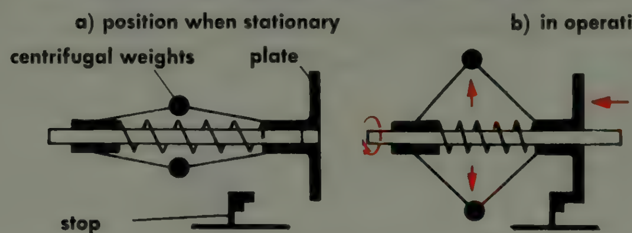


Fig. 5 SPEED GOVERNOR

The cine projector (movie projector) operates on the same principle as the cine camera (see page 214). But whereas in the camera the light rays from the subject are concentrated on the film by the lens, in the projector the rays from a source of light pass through the film image and are projected by the lens on to a screen.

The construction of a cine projector is shown schematically in Fig. 1. Each image on the film is projected on to the screen for a short time, and the film then moves on so as to bring the next image into position for projection. During this movement the film gate of the projector is obscured. To ensure that there is always sufficient slack in the film—because the feed sprocket rotates at a constant rate, whereas the feeding mechanism operates intermittently—a loop of film must be formed between the sprocket and the film gate.

The feeding mechanism which moves the film along may be in the form of a claw (Fig. 2). The lower end of the claw is eccentrically mounted on a disc which rotates at a uniform speed. In addition, the claw is pivoted at the centre. As a result, the head performs the motion indicated in Fig. 2: the points of the claw engage with the perforations in the film and, with each movement, pull it down a distance corresponding exactly to one image. During the time when the image is being projected, the film is stationary and the head of the feeding claw returns to its initial position. A different type of feeding mechanism embodies a so-called maltese cross (Fig. 3). This device is mounted just above the lower film loop and is driven by a constantly rotating disc provided with an eccentric pin. The maltese cross in turn drives a sprocket whose teeth engage with the perforations of the film. Fig. 3a shows the cross in the stationary position. The eccentric pin then engages with the slot in the cross and rotates it with a jerking movement which is imparted to the sprocket (Fig. 3b).

During the time when the film is being moved along, the film gate must be obscured. This is done by the rotary disc shutter, whose mode of functioning is illustrated in Fig. 4. As a result of the intermittent covering and uncovering of the film by this revolving shutter, a flicker effect is produced on the screen. This effect is more pronounced according as the projector light beam is obscured less frequently. To reduce flicker, the image which is being projected is also obscured for part of the time that it is in position at the film gate. For this purpose a three-blade shutter is used, especially with the smaller film sizes.

The film coming from the feeding mechanism forms a loop and then passes over the take-up sprocket, which ensures that, despite the pull exerted by the take-up reel, the lower film loop always remains the same size.

The path of the light rays in the projector is shown in Figs. 1 and 5. The largest possible amount of the light emitted by the projector lamp must be utilised for projection. The condenser lens concentrates the light which is emitted in the forward direction, while the collecting mirror behind the lamp reflects the rearward emitted light. This concave mirror is so mounted that the filament of the lamp is located at the centre of curvature of the mirror. By this means uniform illumination of the film image is obtained. Projector lenses are usually of fairly simple construction (large focal length, small angular field). Projectors are driven by electric motors. A fan or blower for cooling the lamp and film guide equipment is usually mounted on the drive shaft of the motor.

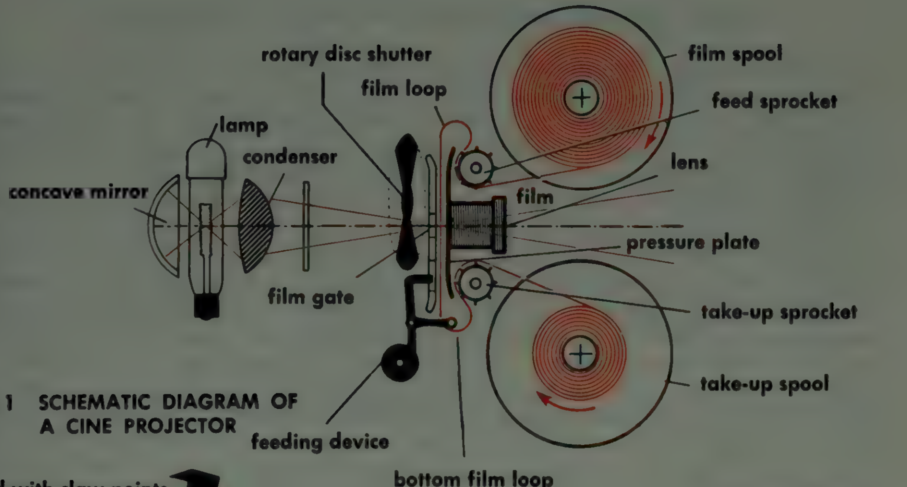


Fig. 1 SCHEMATIC DIAGRAM OF A CINE PROJECTOR

Fig. 2 CLAW TYPE FEEDING MECHANISM

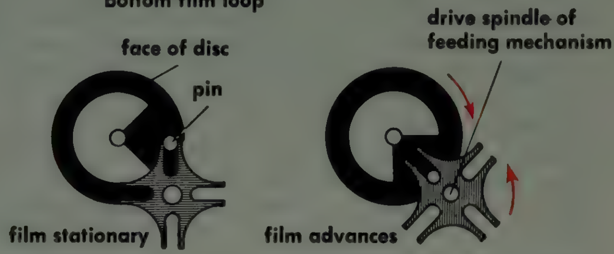


Fig. 3 MALTESE CROSS FEEDING MECHANISM

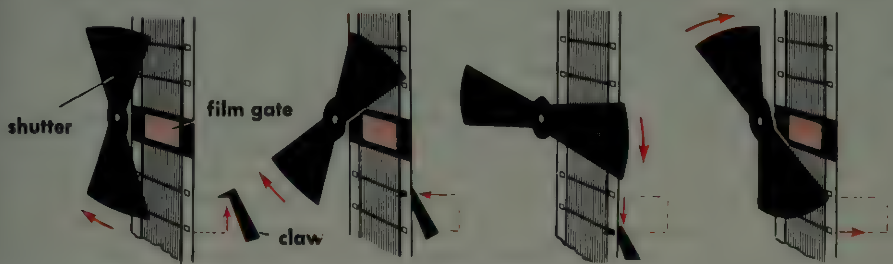


Fig. 4 OPERATION OF ROTARY DISC SHUTTER

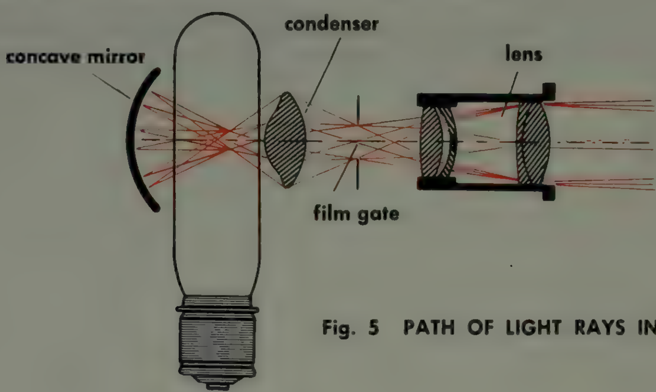
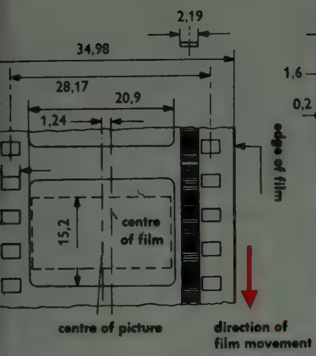


Fig. 5 PATH OF LIGHT RAYS IN PROJECTOR

The filming and projection system described in this article arose from the fact that the "aspect ratio" (width-to-height ratio of the picture) of the conventional cinema does not correspond at all well to the normal field of vision of the human eye. In addition, it was felt to be necessary to improve the quality of the sound, especially with regard to its directional effect (stereophonic sound).

CinemaScope is the proprietary name of a filming and projection system which, by means of special optical equipment, enables cinematographic pictures of great width to be reproduced on a wide screen. The system operates with single-track optical sound reproduction or, better, with multi-channel magnetic sound reproduction in which the source of the sound (loudspeakers installed behind the screen) corresponds to the visual image. CinemaScope film and standard film are both 35 mm wide. When sound film was first introduced in cinematography the frame lines (the strips between the successive frames or pictures on the film) had been substantially widened in order to retain the same aspect ratio of the picture, which had had to be made narrower in order to accommodate the sound track on the film. The CinemaScope picture utilises this otherwise lost space and reverts to the image height that was formerly employed in the old silent film days. The four sound tracks of the CinemaScope system with magnetic sound reproduction are necessary for obtaining "stereophonic" sound with lifelike qualities and proper spatial relationships. The sound is recorded on magnetic tape by three microphones installed in different positions, so that the sounds picked up by each microphone differ in intensity and timing (because of the varying distance to the source of the sound) from those picked up by the other microphones. When the film is subsequently shown, the sound is correspondingly reproduced by three sets of loudspeakers installed on the left and right and at the centre, behind the screen. The fourth sound track is used for sound effects which are not represented as coming from the actual scene projected on the screen and which are reproduced by speakers installed in various parts of the auditorium. The camera for taking CinemaScope films is equipped with an optical attachment called an anamorphotic system, which has the function of squeezing the image together sideways, i.e., the image is distorted by reducing the horizontal scale in relation to the vertical scale. When the film is projected, the process is reversed: the projector is fitted with an anamorphotic system which "stretches" the picture so as to obtain the correct proportions on the screen. Anamorphotic systems may be composed of lenses (Fig. 2), mirrors (Fig. 3) or prisms (Fig. 4). In Fig. 2 the two cylindrical lenses of the anamorphotic system are mounted in front of the projection lens and produce their optical effect by expanding the picture in the horizontal direction (expansion factor 2). An anamorphotic mirror system functions in much the same way as an anamorphotic lens system, but involves a certain vertical displacement of the beam of light. Anamorphotic prisms have the advantage that the magnification of the system can be varied by rotating the prisms.

For physiological and geometrical reasons the CinemaScope screen should be curved. This curvature is particularly essential if the screen is of the reflecting type (cf. page 220).



. 1a DIMENSIONS OF STANDARD FILM (*in mm*)

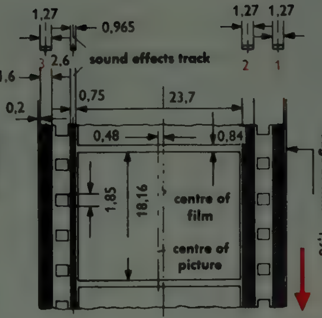


Fig. 1b DIMENSIONS OF CINEMASCOPE FILM WITH FOUR MAGNETIC SOUND TRACKS

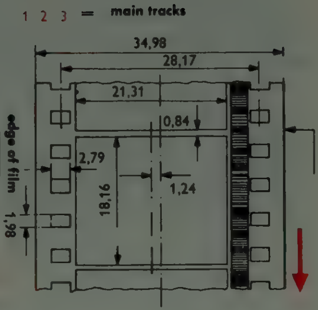


Fig. 1c DIMENSIONS OF CINEMASCOPE FILM WITH OPTICAL SOUND TRACK

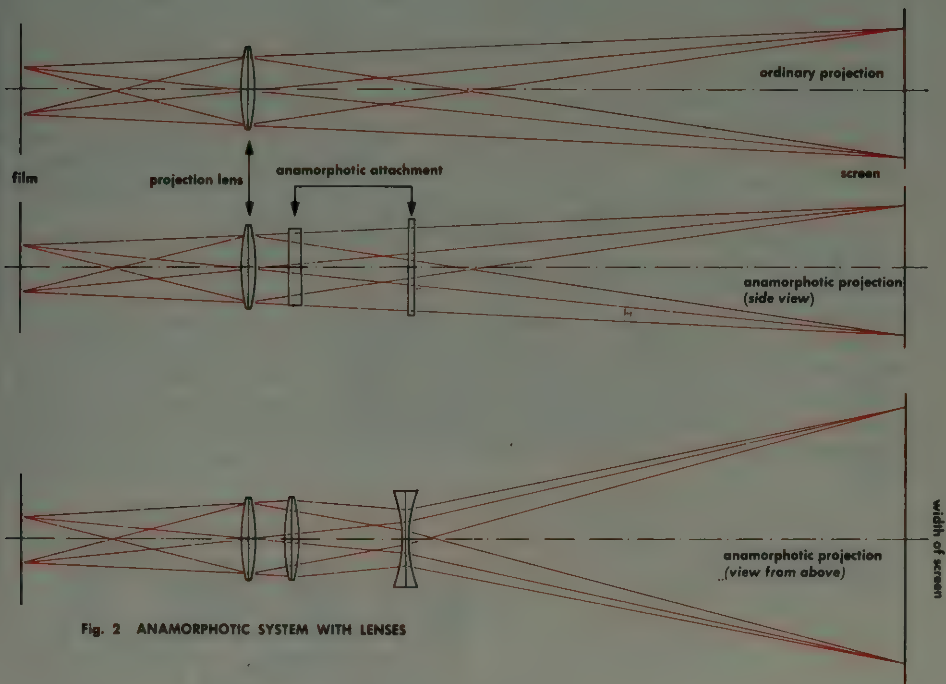
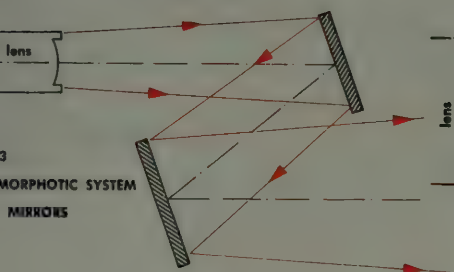


Fig. 2 ANAMORPHOTIC SYSTEM WITH LENSES



3
MORPHOTIC SYSTEM
MIRRORS

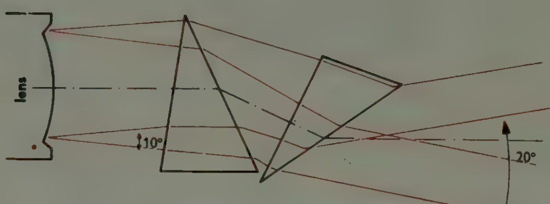


Fig. 4 ANAMORPHOTIC SYSTEM WITH PRISMS

Another modern wide-screen system is Todd-AO. Like the Cinerama system described below, the object is to give the audience a sense of personal participation in the action of the film. The Todd-AO system uses a film which is about twice as wide as standard film, i.e., about 70 mm (Figs. 1a and 1b). Because of the larger picture area, only a moderate amount of enlargement in projection is possible, and this makes for better picture definition on the screen. The sound is recorded on six magnetic sound tracks by means of six microphones. Five tracks correspond to the various directions of the scene, while the sixth track is used for sound effects (cf. page 218). The camera used for filming in Todd-AO is of special design, because of the wide film used and the effect to be achieved: wide-angle lenses are employed (shooting angles of up to 128°; cf. page 180). The most striking feature of the system is the greatly curved wide screen. The audience's sense of direct participation in the scenes shown is very largely attributable to this curvature of the screen. This is bound up with the physiological and psychological processes of visual perception.

Objects projected near, say, the extreme left-hand edge of a flat wide screen will appear considerably compressed in the lateral direction to an observer seated on the extreme right in the front part of the auditorium (Fig. 2a). This distortion is partly corrected by the considerable curvature of the screen. The same is true of the distortions due to the projection itself. The curvature of the screen also avoids or reduces the amount of accommodation to varying distances that the observer's eye has to perform when viewing the picture as a whole. If the screen were made of a white, diffusely reflecting material, the various parts of a curved screen would reflect light towards one another, resulting in a considerable loss of picture definition (Fig. 2b). This disadvantage can most be overcome by using a screen whose surface is provided with a large number of vertical ridges (formed of white plastic), as shown in cross-section in Fig. 2c: the light which strikes any particular point on such a screen is almost entirely reflected towards the audience.

As distinct from the Todd-AO system, which uses only a single camera and a single projector, the Cinerama system makes use of three cameras and three projectors for 35 mm film. The lenses of the triple camera cover a field of about 146°, which corresponds approximately to the field of vision of the human eye. In the cinema the three separate but synchronised films are projected simultaneously on to a semicircular panoramic screen (Fig. 3). The sound system of Cinerama is similar to that of Todd-AO, i.e., six magnetic sound tracks comprising five main tracks and one effects track, except that the sound is reproduced by running a separate "blind" film, with sound recording only, through a special fourth projector.

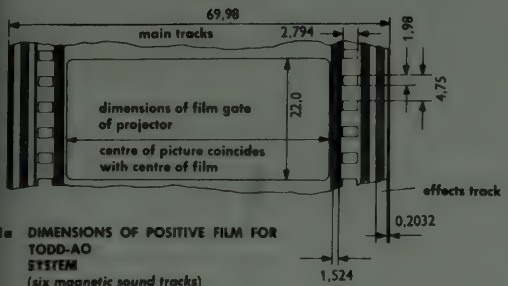


Fig. 1a DIMENSIONS OF POSITIVE FILM FOR TODD-AO SYSTEM
(six magnetic sound tracks)

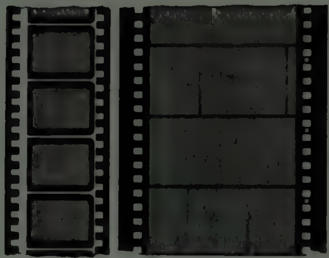


Fig. 1b STANDARD FILM COMPARED WITH TODD-AO COPY

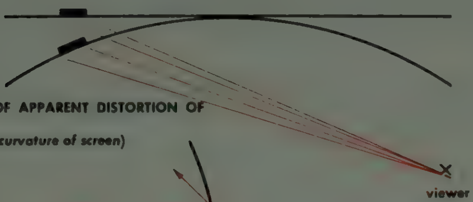


Fig. 2a REDUCTION OF APPARENT DISTORTION OF PICTURE
(correction by curvature of screen)

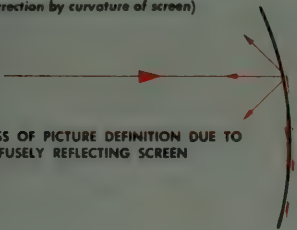


Fig. 2b LOSS OF PICTURE DEFINITION DUE TO DIFFUSELY REFLECTING SCREEN

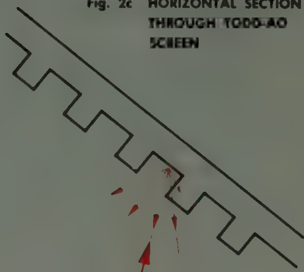


Fig. 2c HORIZONTAL SECTION THROUGH TODD-AO SCREEN

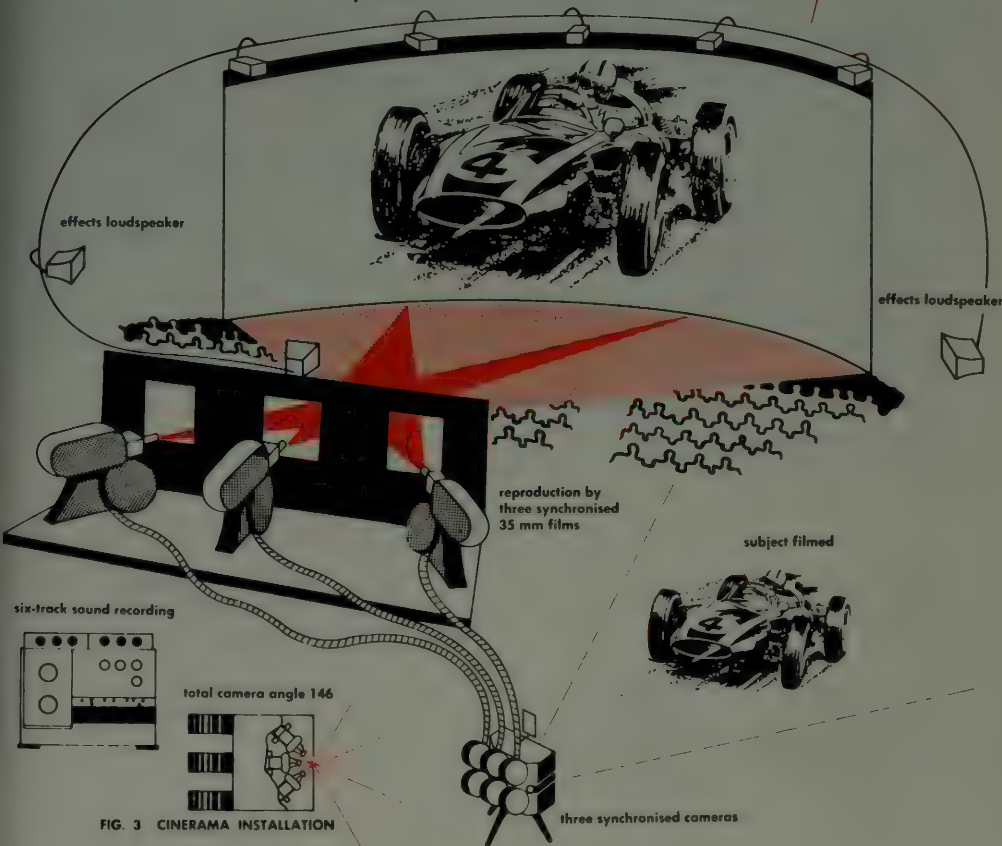
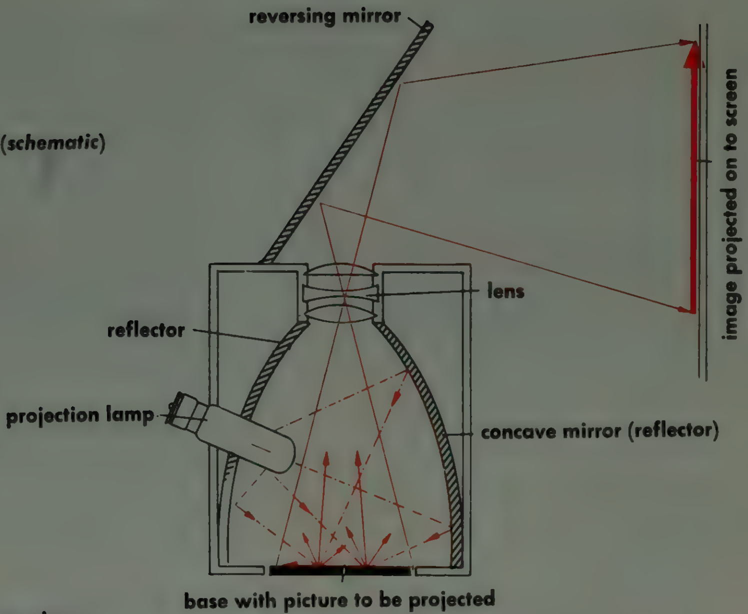


FIG. 3 CINERAMA INSTALLATION

The *episcope* is a projector for opaque flat pictures (paper prints, pages of books, etc.). Its mode of operation is shown schematically in Fig. 1. The efficiency (light yield) of the episcope, i.e., the ratio of the brightness of the image on the screen to the power output of the projection lamp, is relatively poor. This is due to the multiple reflections that some of the light has to undergo inside the apparatus. The filament of the lamp sends rays out in all directions. Only a small proportion of these rays directly impinges upon the picture to be projected. The larger proportion is directed on to the picture by reflectors. From each individual point of the picture rays are sent out in all directions, and only a fraction of these rays passes through the lens and is projected on to the screen via the reversing mirror. A large proportion of the rays is thrown back on to the picture by the reflectors. The light efficiency can be improved by skilful arrangement of the reflectors and by the use of wide-angle lenses of high light-transmitting power. The instrument illustrated in Fig. 2 is known as an "epidiascope"; it can alternatively be used for the projection of transparent slides.

The *diascope*, or slide projector, is intended for the projection of transparencies (diapositives, negatives). Its principal parts the light source with a concave collecting mirror (Fig. 3), the condenser and the projection lens. The source of light (usually an incandescent lamp with a coiled filament) emits light in all directions. The light which strikes the concave mirror is reflected and focused in the plane of the filament. The light emitted in the forward direction encounters the condenser, which in the larger projectors comprises two plano-convex lenses (double condenser). It collects the rays over a large angle and forms the image of the filament at a point located within the projection lens system. The condenser ensures uniformly bright illumination of the entire area of the slide. The light efficiency of a projector is higher according as the angle of the beam of light admitted by the condenser is larger. This angle can, for instance, be increased by placing the condenser closer to the lamp. The refractive power of the condenser will then have to be greater, since the distance from the lamp to the projection lens must remain the same. This requirement can be fulfilled by means of a meniscus lens which is interposed between the double condenser and the lamp (Fig. 5). The efficiency is thereby increased; the image of the lamp filament within the projection lens system is larger. However, because of the shorter distance between lamp and condenser, the amount of heat from the lamp that is absorbed by the condenser becomes greater. The condenser lenses consequently undergo thermal expansion and must be so mounted that this expansion can freely take place. Some of the heat is transmitted through the condenser lenses, however. To protect the slide, a plate of special heat-absorbing glass is installed inside the condenser. This plate itself becomes very hot, and for this reason it is often composed of a number of strips in order to prevent the internal thermal stresses becoming too high and fracturing the glass. Fig. 4 shows a section through the epidiascope functioning as a diascope.

Fig. 1 EPISCOPE (*schematic*)



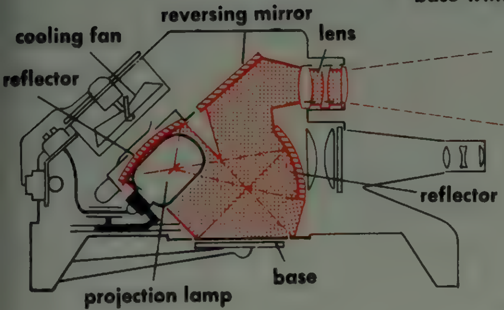


Fig. 2 EPIDIASCOPE FUNCTIONING AS EPISCOPE

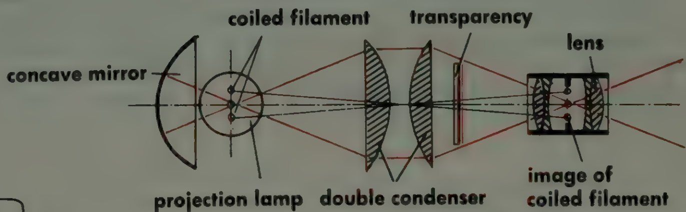


Fig. 3 DIASCOPE WITH DOUBLE CONDENSER (*schematic*)

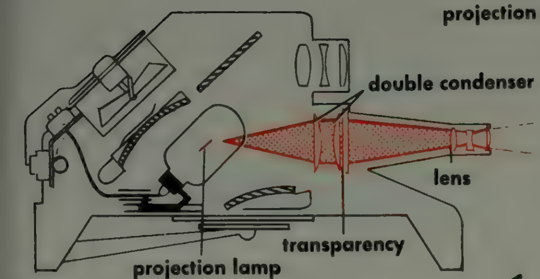
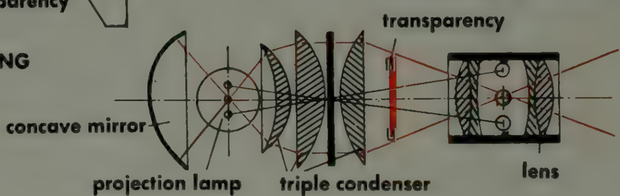


Fig. 4 EPIDIASCOPE FUNCTIONING AS DIASCOPE

Fig. 5 DIASCOPE WITH TRIPLE CONDENSER (*schematic*)



Ordinary photography is based on the light-sensitive properties of silver halides, and more particularly silver bromide (see page 226). The image is developed by chemical processes performed in the liquid phase. Electrophotography, on the other hand, is based on photo-electric and electrostatic effects, and the image is developed by a dry process. The light-sensitive layer in this case is formed by the surface of a photo-semiconductor (cf. page 120). Such semiconductors have a very high dark resistance (up to 10^{14} ohm), whereas exposure to light greatly lowers the resistance (by a factor ranging from 10^5 to 10^7). Materials possessing this property are, for example, selenium (Se), cadmium sulphide (CdS), and zinc oxide (ZnO). Thin selenium coatings applied to an earthed metal base have proved particularly suitable for electrophotography. Sensitisation is achieved by spraying positively charged ions from a corona discharge on to the surface. To this end, a network of thin parallel wires is suspended a short distance from the surface. The wires are given a potential of about + 8000 volts (Fig. 1a). As a result of this, the surface of the coating is charged to about + 600 volts in relation to its back, where a corresponding negative inductive charge is produced (Fig. 1b). On exposure to light, a discharge occurs in those parts of the coating where the light impinges and thus lowers the resistance, so that an equalisation of electric charge between front and back of the coating occurs. In Fig. 2a it is assumed that light impinges on the left-hand and right-hand side of the semiconductor, but that the centre remains unexposed to light. The "image" is formed latently as a pattern of electric charges of varying magnitude. It is "developed" by means of a very fine-grained electrically charged powder (e.g., coloured plastic), which is called the toner. The movement of these powder particles (approx. $\frac{1}{1000}$ mm diameter) is represented in Fig. 2a. Depending on the electric charge of the toner particles, it is possible to bring out and make visible a positive or a negative portion of the invisible pattern of charges. By electrical means this pattern can be transferred to insulating surfaces (e.g., paper) as often as may be desired and can be developed there. Fixation is effected by heating, whereby the particles of plastic are melted, so that they no longer need the electrostatic forces in order to remain adhering to the surface of the paper. The electrical conductivity is produced by means of photo-ionisation due to a photo-electric effect (Fig. 2b): a light quantum (cf. page 120) dislodges an electron from an atom and thereby causes the formation of a pair of electrically charged particles. The negative electron neutralises a positive elementary charge at the surface; the positively charged atomic nucleus travels to the back and neutralises a negative charge there. The formation of a visible image respectively in the negative-positive process and in the positive-positive process is illustrated diagrammatically in Fig. 3a and Fig. 3b.

Fig. 1a

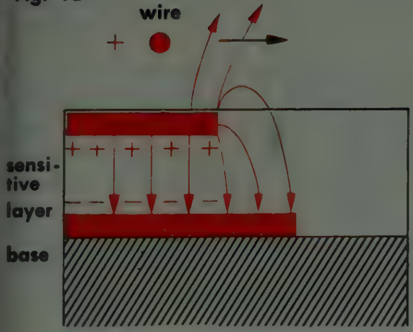


Fig. 1b

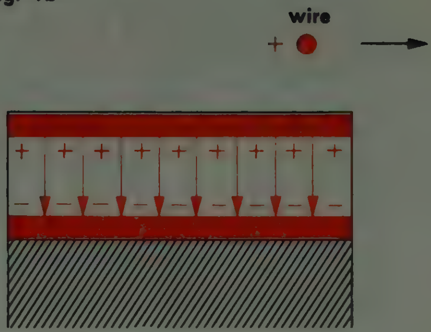


Fig. 2a

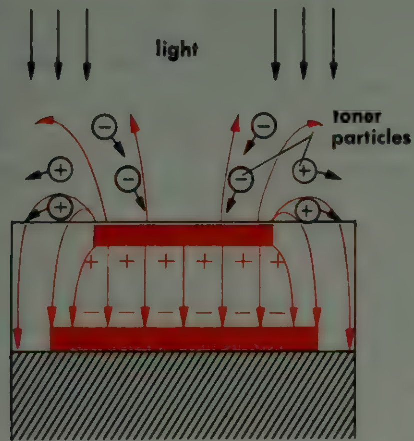


Fig. 2b

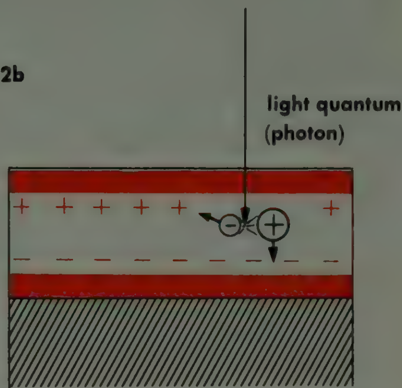


Fig. 3a

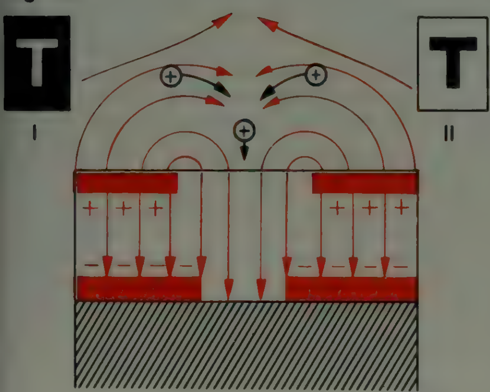
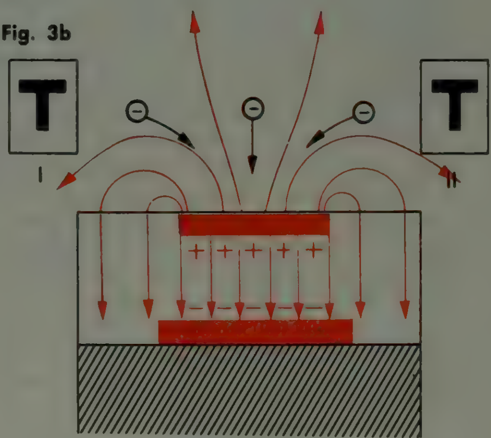


Fig. 3b



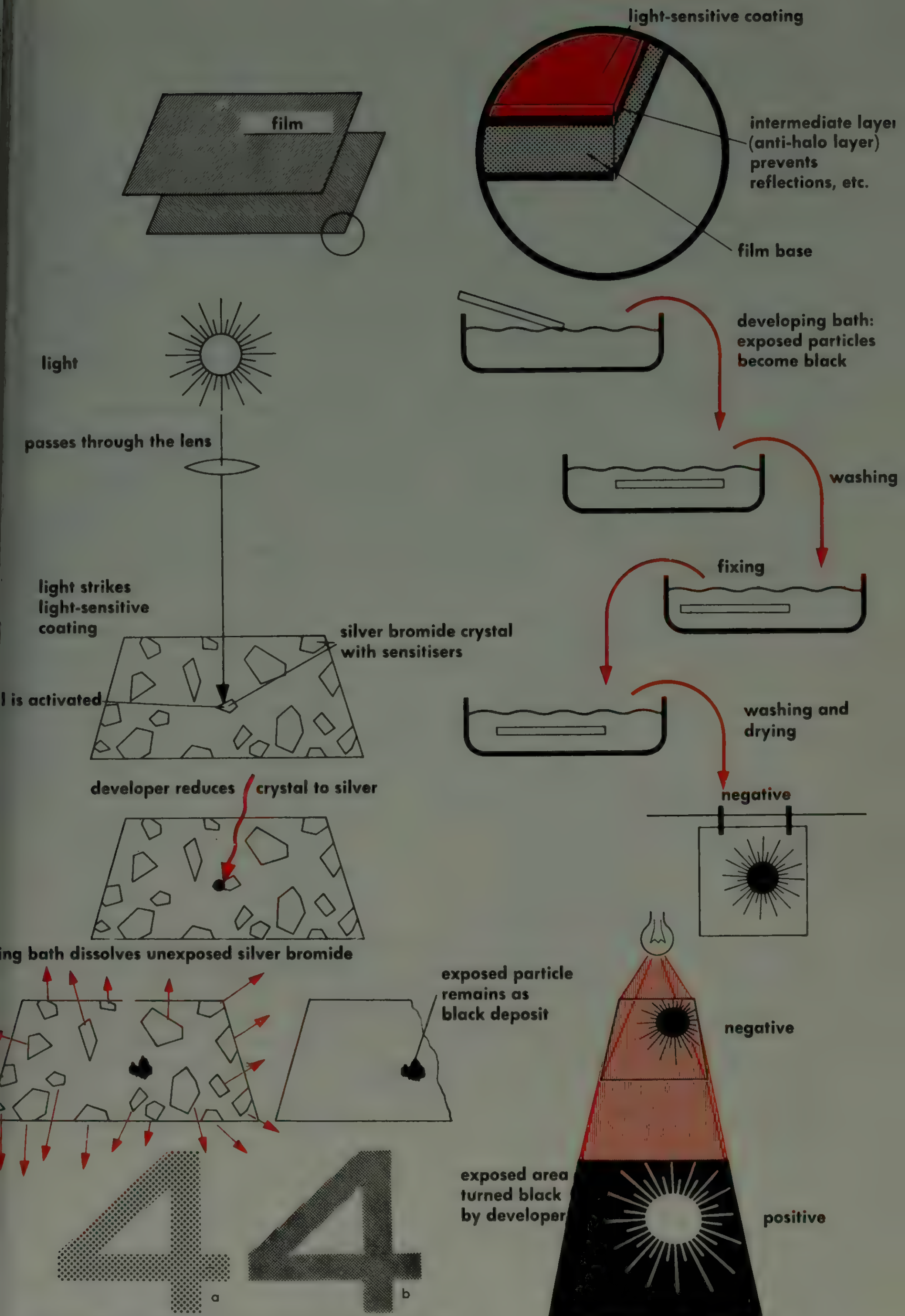
I = pattern of electric charges II = toner image

Black-and-white film is a highly light-sensitive photographic material which is sold commercially in the form of roll films, plates or photographic papers. It is used for black-and-white photography.

The film usually consists of a very thin base made of cellulose acetate, nitrocellulose (cellulose nitrate) or metal. The light-sensitive emulsion is applied to a thin dark-coloured intermediate layer (anti-halo layer). The emulsion is in the form of a very thin coating consisting of 40% silver bromide crystals, 50% gelatine (as a binder) and 10% water. In complete darkness the silver bromide is emulsified in gelatine solution and, together with further additives, is applied by machine to the film or plate. The size of the tiny silver bromide crystals determines the light-sensitivity ("speed") and the resolving power of the emulsion coating. A coarse particle is highly light-sensitive; on the other hand, its optical resolving power is low. Pictures taken with coarse-grained film cannot be greatly enlarged, as the enlargement obtained would be too "grainy". The silver bromide emulsion is not equally sensitive to all colours of light. Yellow-green, yellow, orange and red colour shades could not be photographed on such film. For this reason the coating is chemically sensitised by the addition of minute quantities of gold, mercury and other heavy metal ions to the silver bromide and by feeble reaction with sulphide ions. Also, the coating is physically sensitised by the addition of pigmental sensitisers and thus made sensitive to a wider range of hues. Orthochromatic film is sensitive to the blue-green-yellow range of light, orthopanchromatic film is sensitive to these colours and additionally extends the range to orange, while panchromatic film is sensitive to the whole range of visible light including red. There are, furthermore, special kinds of photographic film with widely varying sensitivity properties for scientific research purposes.

The photographic image is formed by the light rays which come from the subject (i.e., the object being photographed) and are focused on the light-sensitive coating of the film by the lens system of the camera. Each ray of light that strikes the film encounters silver bromide crystals, which are thereby activated. The exposed film is then developed in a development bath, which is a solution of a chemical reducing agent (developer). The latter reduces the activated silver bromide particles to black metallic silver. Silver bromide which has not been activated by exposure to light is not reduced by the developer and can be dissolved out of the emulsion coating by immersion of the developed film in a fixing bath, which contains a solvent (usually solium hyposulphite or "hypo") that dissolves unexposed silver bromide.

After the developer and fixing solutions have been washed out of the film and the latter has dried, the result obtained is the negative. (What has been described here is the simplest procedure; for special purposes the negative can be subjected to various intermediate chemical and heat treatments, etc.) On a negative the shadows appear as "white" (transparent) areas, while the highlights appear black. If the negative is placed in contact with photographic paper (positive paper) and exposed to light, then the areas of this paper under the (dark) highlights in the negative will receive little light, while the areas under the (white) shadows in the negative will be strongly exposed. On developing this photographic paper, a "print" is obtained in which the exposed areas come out dark, while the unexposed areas remain white, and of course those areas which are somewhere between deep shadow and bright highlight come out in various intermediate shades of grey. The dark and light areas of the negative are thus reversed in the printing process.



black silver particles: coarse grain (a), fine grain (b)

By noise is understood sound consisting of a mixture of airborne vibrations which is completely irregular with regard to sound intensity, frequency and phase. Noise is usually regarded as a nuisance and may, if the sound intensities involved are very high, cause damage to the organs of hearing. Such objectionable noise is, for example, caused when a piece of sheet metal is struck with a hammer (Fig. 1a). When the sound intensity is plotted as a function of time, the curve obtained presents a very irregular shape with jagged peaks and valleys (Fig. 1b). To provide protection against noise is an important function of modern structural engineering. For example, in offices in which the rattle of typewriters would produce unbearable conditions if it were reflected back from the walls, the ceilings and, if need be, the walls can be lined with sound-absorbing panels (Fig. 2). The actual sound-absorbing material, e.g., glass wool or rock wool, is interposed between the ceiling (or wall) itself and an inner skin formed of perforated plates secured at a certain distance from the wall face. The sound waves which pass through the perforations are absorbed by the glass wool or similar absorbing material. The result is a remarkable reduction of noise. According to the Weber-Fechner law of hearing, the apparent loudness of a sound—i.e., the intensity of the acoustic perception—is approximately proportional to the logarithm of the intensity of the stimulus (sound intensity). The unit of objective loudness or sound level is the phon. The loudness, in phons, of a sound is equal to the intensity in decibels of a sound of frequency 1000 cycles/sec. which seems as loud to the ear as the given sound. Two sound intensities, P_1 and P_2, are said to differ by n decibels when $n = 10 \log_{10} P_2/P_1$, where P_2 is the intensity of the sound under consideration and P_1 is the intensity of the reference sound level.

The following table gives typical loudness values, in phons, for a number of sound sources:

air raid siren	135 phons
aircraft at take-off	130 phons
pneumatic hammer	120 phons
hooter	110 phons
engineering shop	100 phons
underground railway train	95 phons
heavy lorry (truck)	90 phons
motor cycle	85 phons
moped (motor bike)	80 phons
office	75 phons
road traffic	70 phons
transformer	60 phons
rustling leaves	30 phons
soundproofed room	10 phons

Loudness values in excess of 130 phons produce a sensation of acute discomfort. Continuous noise in excess of 100 phons may cause hearing damage. The sound-level scale from 1 to 140 phons comprises the intensity ratio of $1:10^{14}$. Objective measurement of sound intensity can be carried out by means of a Rayleigh disc (Fig. 4). This is a small thin disc which is suspended from a fine thread of glass or quartz and is placed at an angle of 45° to the direction of propagation of the sound waves. The disc experiences a torque and strives to place itself at right angles to the waves. The amount of rotation it undergoes is a measure of the intensity. See also page 242, vol. II.

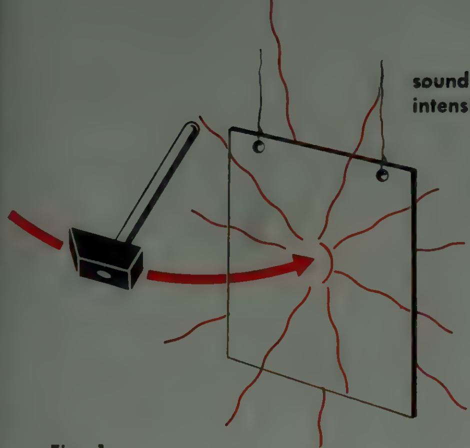


Fig. 1a

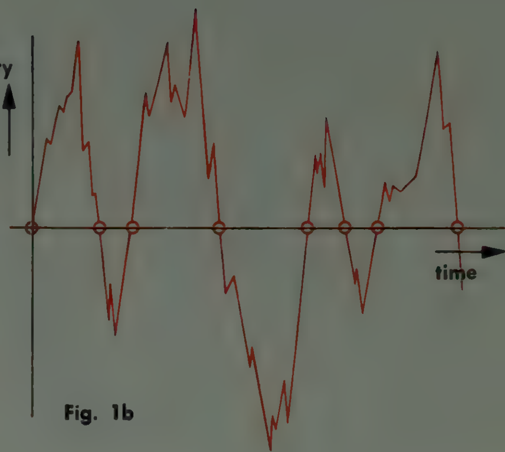


Fig. 1b

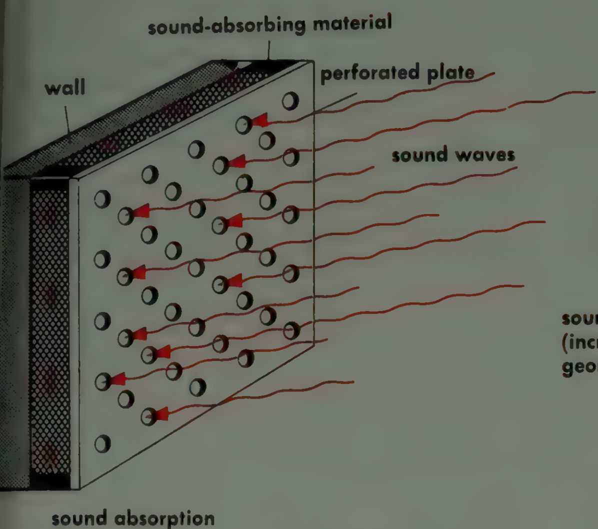


Fig. 2

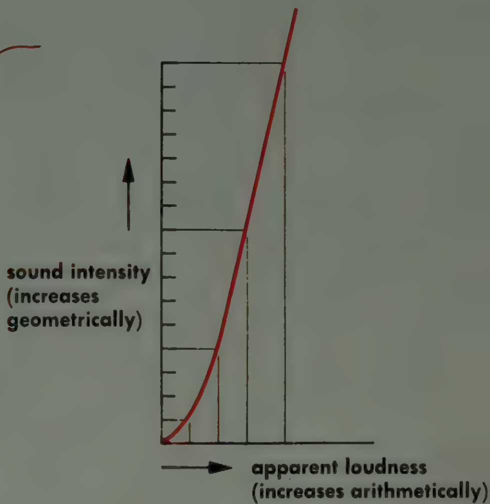


Fig. 3 DIAGRAM REPRESENTING THE WEBER-FECHNER LAW OF HEARING

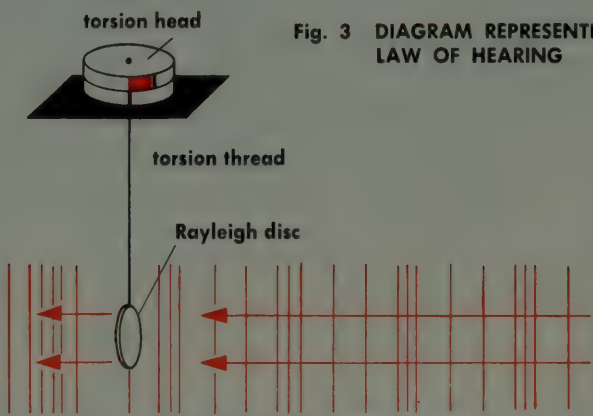


Fig. 4 SOUND INTENSITY MEASUREMENT

Acoustical chamber at Bell Laboratories
Photo USIS

By resonance is understood the phenomenon that structures capable of oscillation will oscillate in sympathy with relatively feeble external forces which act periodically and whose oscillation period coincides with that of the resonating structure. While it is resonating, the structure stores up energy. Under certain circumstances the amount of energy stored up in this way may become so great that it brings about the destruction or collapse of the structure. A simple example of a resonating structure is a child's swing (Fig. 1a). It is a pendulum which is given a push or a thrust in the swinging direction each time it reaches its maximum deflection. Its energy build-up, i.e., its resonance, is directly evident from the increasing amplitude of the deflection of the swing. Another example is a liquid in a U-shaped tube (Fig. 1b). The liquid can be set in motion by blowing into one end of the tube, and by blowing it periodically at the appropriate instant, the amplitude of its oscillations is progressively increased. The oscillations do not, however, go on increasing indefinitely, but are limited by energy losses — in this case more particularly by losses due to friction of the liquid on the wall of the tube. Resonance of a magnetically polarised steel spring can be induced by the fluctuating magnetic field of an electromagnet energised by an alternating current (Fig. 2a). This resonance effect is, for example, utilised in frequency meters. The conception of resonance had its origin in the science of acoustics. Fig. 2b illustrates an acoustic resonator, a device known as Kundt's tube, which is used for measuring the wavelength of sound waves. Projecting into the glass tube is one end of a metal rod which is held gripped in the middle. Longitudinal vibrations are set up in this rod by rubbing it with a cloth sprinkled with powdered rosin. The end of the rod in the tube is provided with a disc which in turn transmits the vibrations to the air in the tube. The effective length of the tube can be varied by means of an adjustable disc at the other end. The vibrations (sound waves) are reflected by this disc, and on suitably adjusting its position, a stationary wave will be produced in the tube, and resonance occurs. This happens when the distance between the two discs is equal to an odd multiple of one-quarter of the wavelength of the sound waves set up in the tube (see also page 240), and vibration nodes and antinodes are formed. These can be indicated by introducing a small quantity of some suitably light powder (e.g., lycopodium powder) into the tube. The powder congregates in a heap at each node. The nodes are thus made "visible", and the distance between them can be measured. The distance between two successive nodes is equal to half the wavelength of the sound waves set up in the tube. Resonance effects are also observed in connection with electromagnetic phenomena. The most well known and important example is the excitation of an electromagnetic oscillatory circuit, comprising a self-inductance L and capacity C, by an alternating voltage (Fig. 3). In the circuit the energy oscillates between its electrical state in the condenser (Fig. 3a) and its magnetic state in the magnetic field of self-induction (Fig. 3b). If the natural period of vibration (and therefore the frequency) of the oscillatory circuit corresponds to that of the alternating voltage, resonance will occur. The circuit will in that case absorb the maximum amount of energy from the source of energy that produces the excitation. Radio transmitters and receivers are tuned with the aid of this resonance effect (cf. p. 86 and vol. II, p. 260). To prevent the energy attaining disastrously high values, resistances are included in the circuit; these cause energy losses in the form of heat.

Another phenomenon that acoustic and electric vibrations have in common is echo, i.e., the reflection of sound waves or electromagnetic waves from obstacles they encounter (Fig. 4). See also page 136.

Fig. 1a

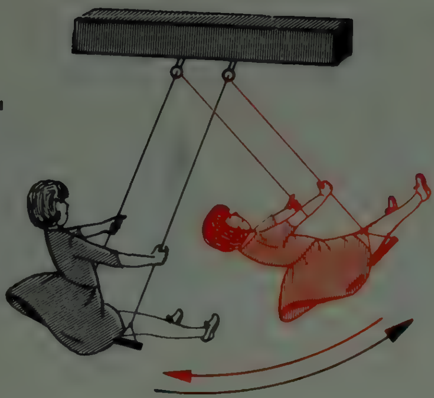
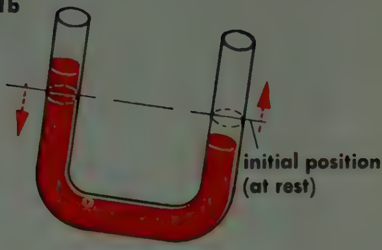
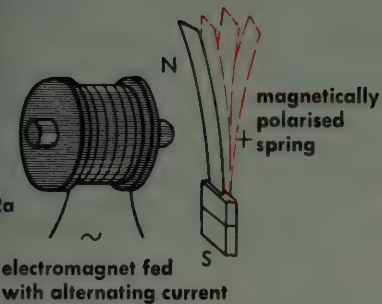


Fig. 1b

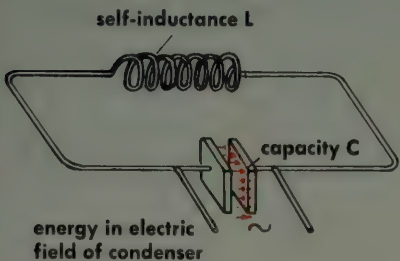




g. 2a

Fig. 2b ACOUSTIC RESONATOR (KUNDT'S TUBE)

a)



b)

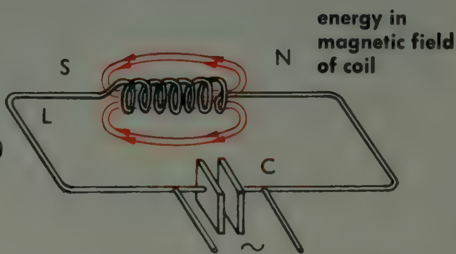


Fig. 3 EXCITATION OF OSCILLATORY CIRCUIT

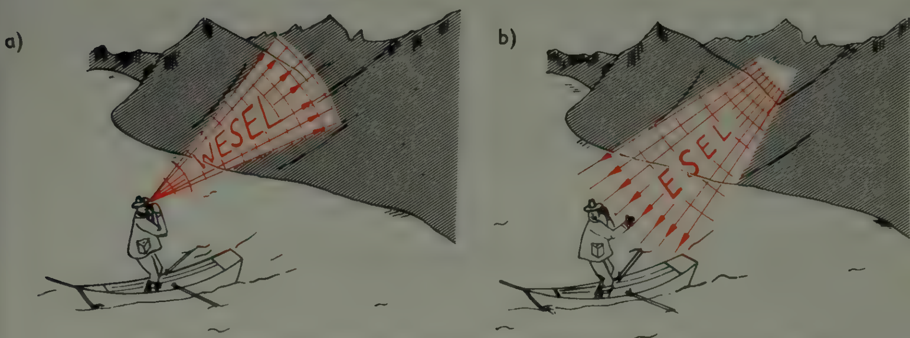


Fig. 4 ECHO IS REFLECTION OF SOUND WAVES FROM AN OBSTACLE

When a vibrating source of waves is approaching an observer, the frequen observed is higher than the frequency emitted by the source. When the source receding, the observed frequency is lower than that emitted. This is known as tl Doppler effect, or Doppler's principle, and is named after an Austrian physici who lived in the first half of the 19th century. Figs. 1 and 2 will help to explain th phenomenon. When a whistling locomotive (or any other sound source) approach a stationary observer (Fig. 1), more density concentrations reach his ear than whe both the sound source and the observer are stationary. As the pitch depends on tl frequency (number of vibrations per second), the sound from the approachin locomotive's whistle has a higher pitch than the sound coming from the same whist when the locomotive is stationary in relation to the observer. Similarly, when tl locomotive is receding, its whistle sounds with a lower note. At the instant when tl locomotive passes the observer, the note of the whistle is heard to change to lower pitch. The same effect is observed when we are passed by a fast-moving hootin car in the street, or when the observer is moving fast in relation to a stationar sound source (e.g., a motor cyclist approaching a siren, as in Fig. 2).

The Doppler effect is widely used in astronomy for measuring the velocity a which distant stars or nebulae are approaching or receding. These motions produc a shift in the position of lines in their spectra (see page 156). A particular spectrur line corresponds to a certain definite light wavelength. If the star emitting the ligh is moving away from us, its light rays have a longer wavelength (lower frequency by virtue of the Doppler principle, and this is manifested in a general shift of th spectrum lines towards the red end of the spectrum. This is known as the "red shift" Similarly, in the spectrum of a star moving towards us, the characteristic lines woul show a "blue shift", i.e., they would be displaced towards the blue end of the spectrum corresponding to shorter wavelengths and higher frequencies. These phenomena ar indicated in Fig. 3. A remarkable thing about the spectra of the spiral nebulae (th galaxies of stars far out in space beyond our own Milky Way system) is that they al display the red shift and must therefore—on the basis of Doppler's principle— all be moving away from us. The theory of the "expanding universe" is based or this phenomenon. However, this interpretation of the red shift is disputed by some authorities.

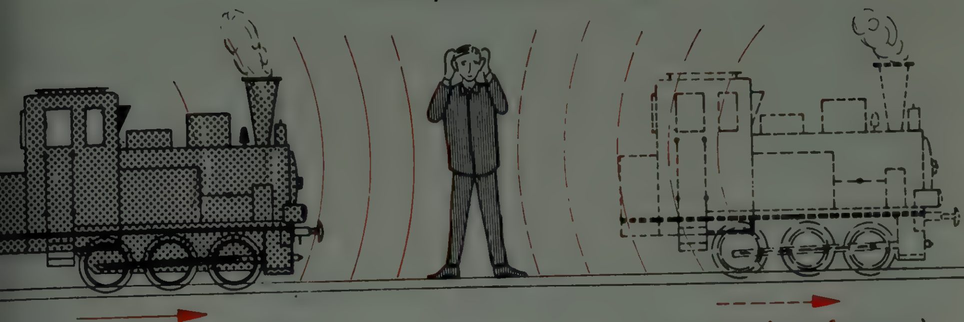


Fig. 1

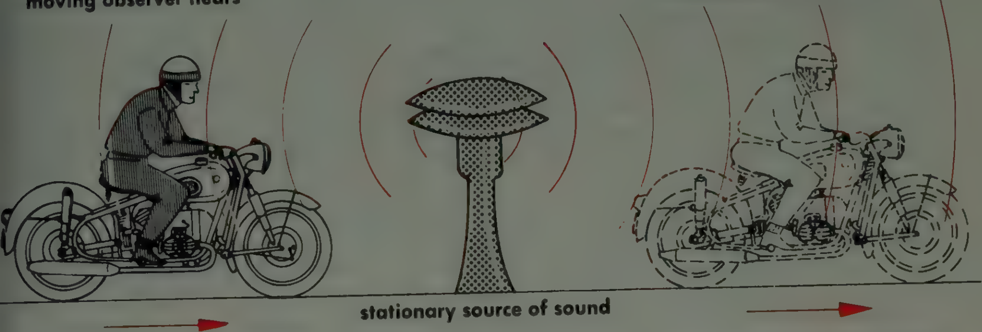


Fig. 2

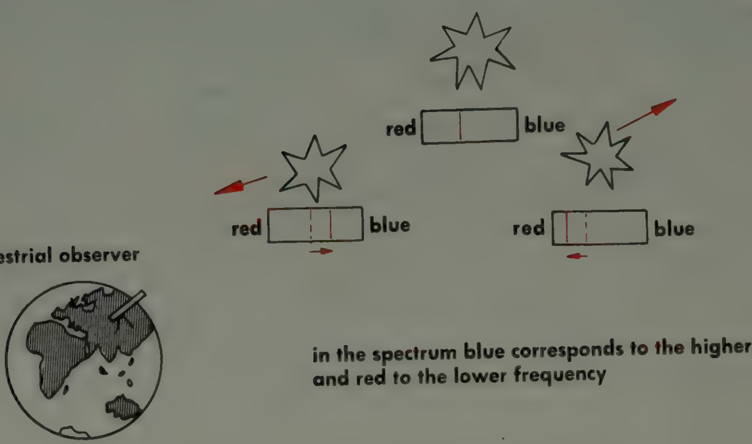


Fig. 3 RED SHIFT INTERPRETED AS DOPPLER EFFECT

The term ultrasonics (or supersonics) refers to sound vibrations—variations of density in elastic media (e.g., air)—whose frequencies are beyond the auditory limit, i.e., above approx. 20,000 cycles/sec. The highest ultrasonic frequencies hitherto attained are of the order of 10 million cycles/sec. Such high-frequency elastic vibrations are produced in various ways, based on different physical principles. An obvious method, in the first place, is to extend the old acoustic principle of sound generation by means of pipes to the ultrasonic range. This can be done by means of the Galton pipe (Fig. 1). This is a pipe (cf. page 240) in which the position of the gap and lip can be varied by micrometer adjustment. It is blown by compressed air and enables ultra-sound with frequencies up to 30,000 cycles/sec. to be produced. Higher frequencies can be attained by making use of other phenomena, namely, magnetostriction and the piezoelectric effect (cf. page 104). Magnetostriction is the change in the dimensions of a ferromagnetic material when it is placed in a magnetic field. If the latter is produced by an alternating current, the material will undergo vibrations. Fig. 2 is a diagrammatic representation of a magnetostriction ultrasonic generator. It consists of 0.1–0.3 mm (0.004–0.012 in.) thick nickel plates, insulated from one another, whereby the eddy current losses are reduced. The general construction is rather like that of a transformer (page 110). The arrangement of the windings, as indicated in the diagram, causes the magnetic field to form a closed circuit in the stack of plates. Utilising this principle, ultrasonic waves with frequencies of up to 200,000 cycles/sec. can be produced. The sound is radiated—e.g., in air or water—in two directions (upward and downward in Fig. 2). If radiation is required in one direction only, one face must be provided with a foam rubber cushion, which acts as a screen impervious to sound.

Even higher frequencies can be attained by means of the piezoelectric effect. A circular quartz plate—cut from a hexagonal quartz crystal in the manner shown in Fig. 3—is so gripped at its edge (Fig. 4) that thin metal foils on the two circular faces of the plate impart a high-frequency charge to it and that it is nevertheless able to vibrate freely. The high-frequency thickness vibrations caused by the piezoelectric effect are transmitted to the air in a sound box and thence via a diaphragm to an adjacent medium (e.g., water).

Ultrasonic vibrations are used in many technical applications, including non-destructive testing of materials, degasification of liquids, echo-sounding, and also in therapeutic medicine. Certain nocturnal animals, such as bats, make use of ultrasonic vibrations to guide their movements in the dark on the radar principle (cf. page 136).

Fig. 1 GALTON PIPE

Fig. 2 MAGNETOSTRICTION ULTRASONIC GENERATOR (*schematic*)

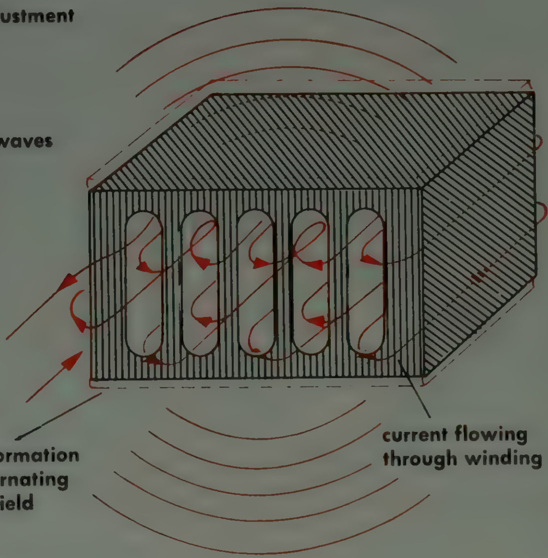


Fig. 3 PLATE CUT FROM QUARTZ CRYSTAL

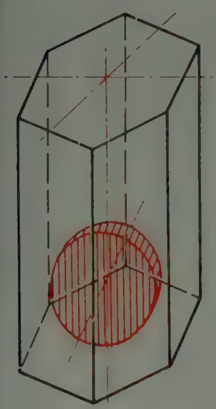
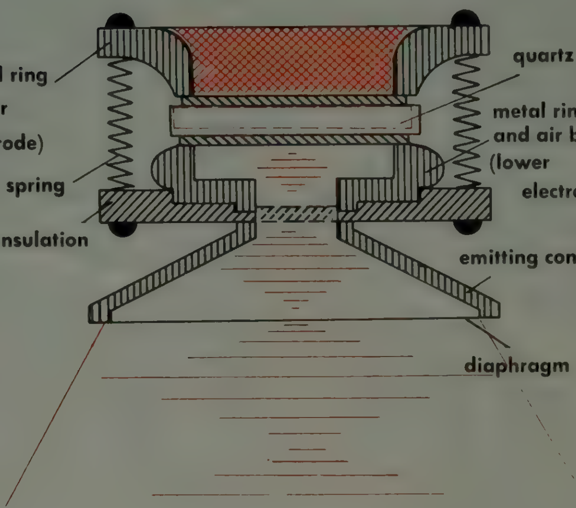
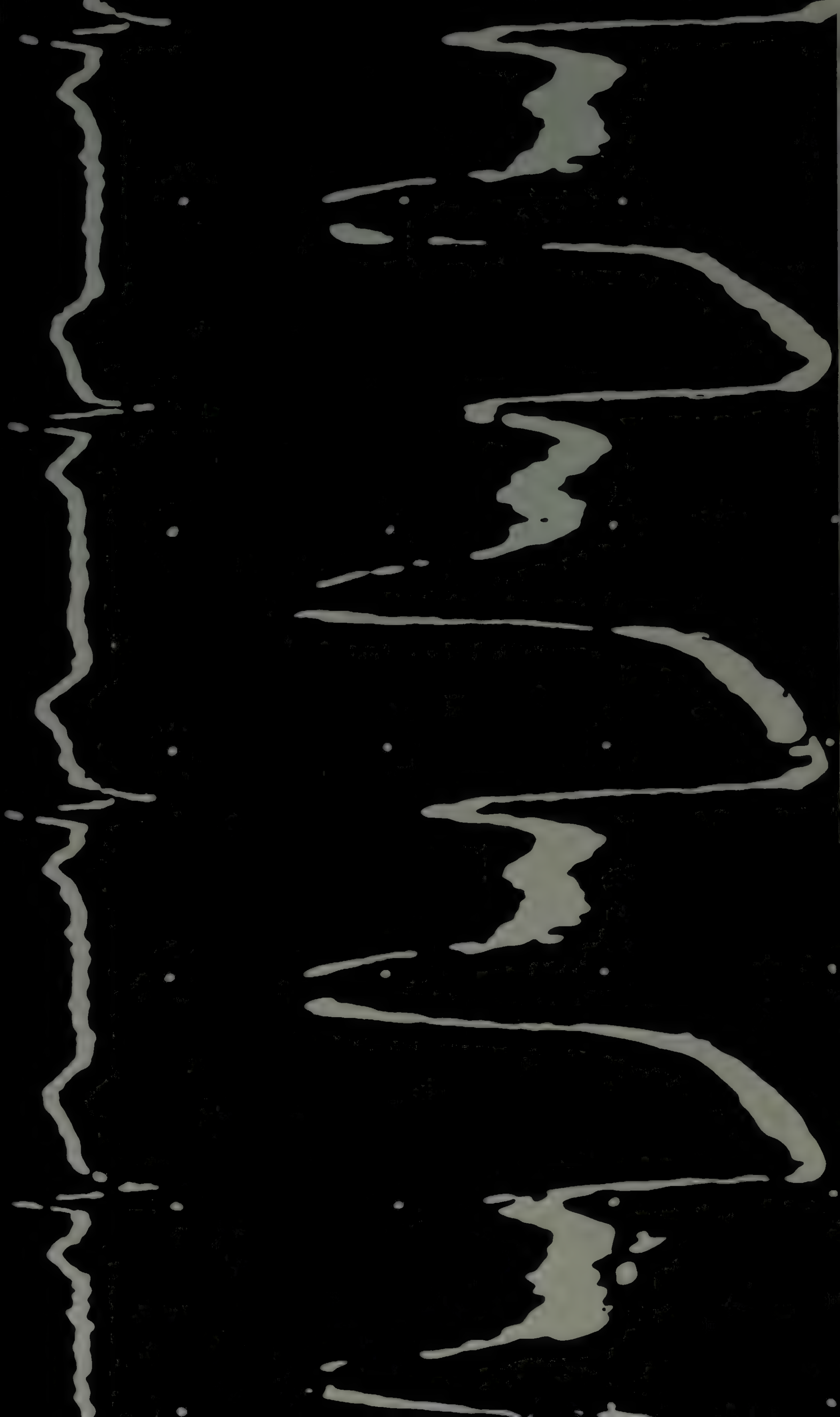


Fig. 4 PIEZOELECTRIC SOUND TRANSMITTER

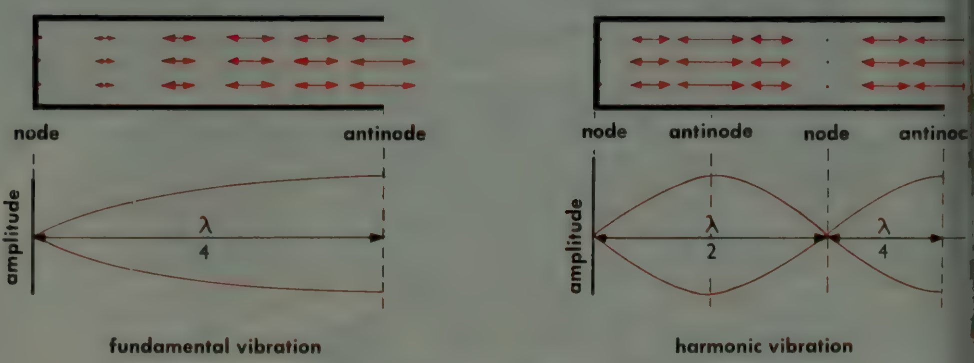


Ultrasonic energy is harnessed to help in detecting diseased tissue cells
Photo USIS



If a vibrating tuning fork is held over a glass cylinder, the column of air in the cylinder will, if it is of appropriate length, resonate with the fork. The length (l) of the air column can be adjusted by pouring water into the cylinder so as to make it an odd multiple (n) of one-quarter of the wavelength (λ) of the note produced by the tuning fork, i.e., $l = n\,\lambda/4$, where n denotes an odd integer; in that case resonance will occur (see page 232). This means that the column of air vibrates "in tune" with the tuning fork, so that the sound emitted by the fork is strengthened. The air in the glass cylinder forms what are known as stationary waves, with nodes and antinodes. At the nodes the air in the cylinder is at rest, whereas it moves to and fro, i.e., it is in a state of vibration, at the antinodes. Fig. 1 shows the arrangement of the nodes and antinodes associated with the second harmonic vibration (the second overtone). The fundamental vibration and the first harmonic vibration in an open-ended pipe or cylinder are indicated in the diagram at the foot of this page. At the antinodes the air periodically increases and decreases in density, in time with the frequency of the vibration.

The operation of all wind instruments for producing sounds is based on resonance vibrations. These vibrations can be generated by two different methods: a stream of air issuing from a gap flows past a sharp-edged lip and produces a vortex motion of the air (labial organ pipe, Fig. 2a; flute, Fig. 3); alternatively, a stream of air is periodically interrupted by means of a vibrating tongue of wood or metal, known as a reed (reed pipe, Fig. 2b). In the organ pipe the air first enters a chamber in which the air is stored up; it then flows through a gap and impinges on a lip which causes vortexes to form. These generate resonance vibrations in the air column in the pipe, whose length (l) is a multiple of half the wavelength, i.e., $l = n\,\lambda/2$, where n is an integer. In Fig. 2 the conditions corresponding to the second harmonic, or overtone, are represented. The flute (Fig. 3) functions on the same principle as the organ pipe. When all the holes are closed, the flute gives its fundamental note, i.e., the note with the lowest frequency and associated with the greatest length of the vibrating column of air. When one of the holes is opened, a higher note is heard according as the vibrating air column is shorter. In the reed pipe (Fig. 2b) the air causes a resilient tongue, called a reed, to close an opening and thus cut off the air flow. As soon as this happens, the reed springs back, and the action is thus repeated periodically, whereby vibrations are set up in the funnel-shaped pipe above the reed. The reed pipe principle is employed in various musical instruments, such as clarinet, oboe, harmonium and mouth-organ.



of a column of air open at one end

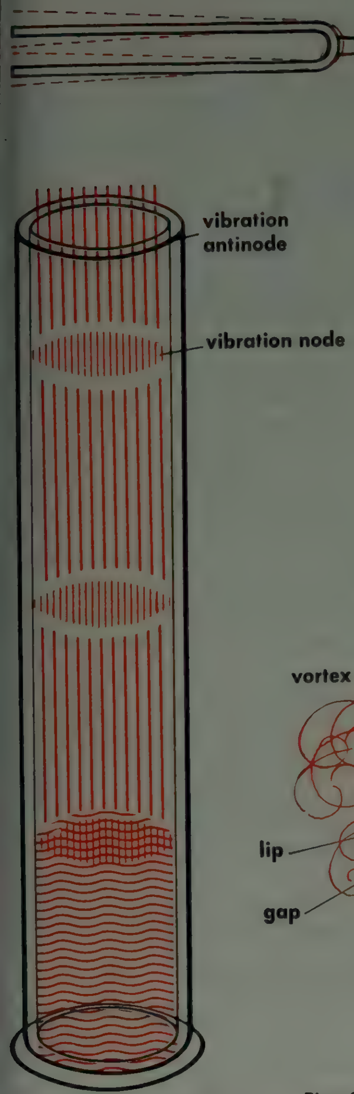


Fig. 1 TUNING FORK PRODUCES RESONANCE VIBRATIONS

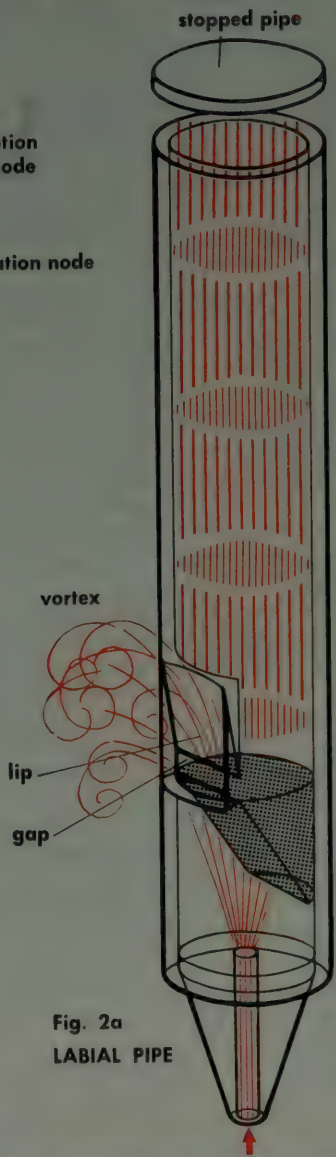


Fig. 2a LABIAL PIPE

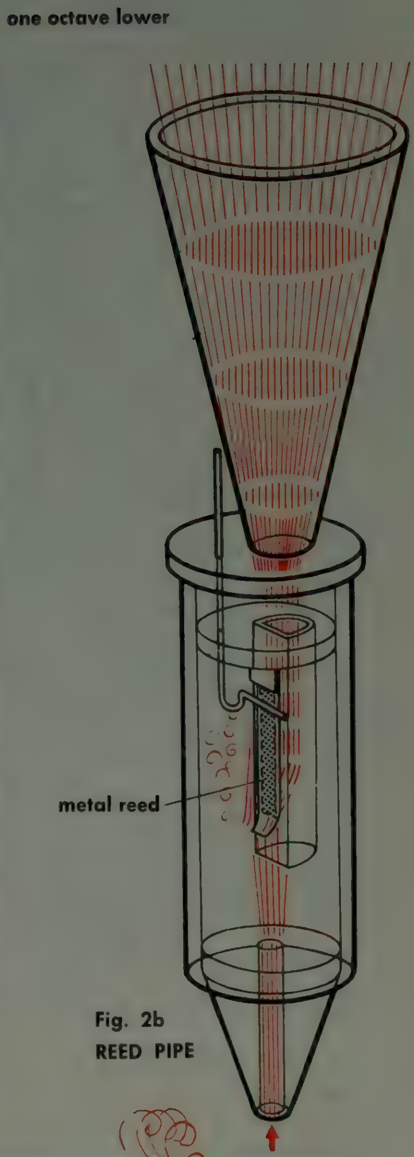


Fig. 2b REED PIPE

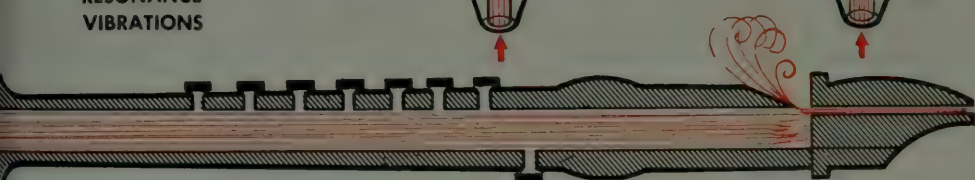


Fig. 3a FLUTE (*fundamental note*)

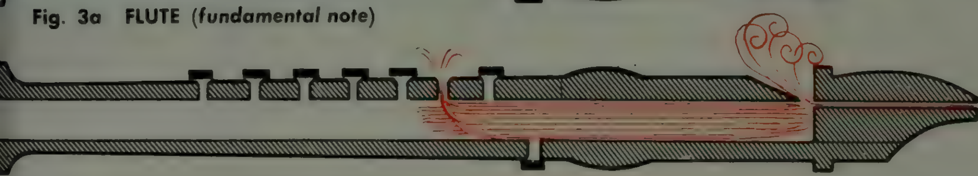


Fig. 3b FLUTE (*note one octave higher than in Fig. 3a*)

The force with which the earth attracts an object, i.e., the gravitational force exerted upon it, is called weight. This force is proportional to the mass (the amount of matter) in the object and to the gravitational acceleration. However, the acceleration of gravity is not exactly the same in all parts of the world: its value is a little higher at the poles than at the equator; in other words, the same object at the poles of the earth has a somewhat greater weight than it has at the equator.

The simplest form of weighing device is the equal-armed beam scale (or balance) (Fig. 1). It consists of a beam pivotably mounted on a knife-edge fulcrum at the centre. Attached to the centre of the beam is a pointer which points vertically downwards when the scale is in equilibrium (Fig. 1a). For proper functioning of the scale it is essential that the centre of gravity of the beam is located lower than the pivot (fulcrum). For this reason the pointer is provided with a vertically slidable weight. The beam is in equilibrium when the clockwise rotating moment (load on right-arm × length of right arm) is equal to the anti-clockwise rotating moment (load on left arm × length of left arm); since the two arms of the beam are equal, the scale will be in equilibrium when the weight of the object to be weighed is equal to that of the weights placed in the other pan. The sensitivity of the beam scale is lower according as the centre of gravity of the beam is located lower down. From Fig. 1b it is apparent that, when the weight on the pointer is lowered, the lever arm of this weight in relation to the pivot will, for the same angle of tilt of the beam, be increased from a to b. The beam will in fact rotate about the pivot until b becomes equal to a, so that the deflection of the pointer decreases, although the loads on the two pans of the scale are the same as before.

The equal-armed scale requires a set of weights at least as heavy as the heaviest load to be weighed. For the weighing of heavier objects, scales having arms of unequal length are therefore used. The weights are suspended from the longer arm, and the object is suspended from the shorter arm. For equilibrium, the forces (loads) acting upon the two arms have to be inversely proportional to the lengths of the arms. Thus, in the weighbridge illustrated in Fig. 2, the weight needed to balance the load is only one-tenth of the latter if the machine is so designed that $l = 10 \times a$ and $a:b = c:d$. The platform weighing machine illustrated in Fig. 3 is of the pendulum type, i.e., it is equipped with a counterbalancing pendulum instead of a pan for weights. The deflection of the pendulum (and therefore of the lever attached to it) provides an indication of the magnitude of the load placed on the weighing platform. To alter the range of the machine, a sliding weight can be moved along a lever.

In the case of the letter balance (Fig. 4) the rods interconnecting the various pivots form a parallelogram in all positions. When a load is placed on the platform, the lever arm of the counterbalancing weight in relation to the pivot increases, so that the deflection provides a measure of the magnitude (i.e., the weight) of the load.

Another type of scale is the spring balance (Fig. 5). In this device the extension of the spring is proportional to the magnitude of the load suspended from it. The scale can be directly graduated in weight units. The familiar "bathroom scales" are also based on the spring balance principle. Here, too, the extension or deflection of a spring is transmitted to a scale whose markings give a direct reading of the weight (Fig. 6).

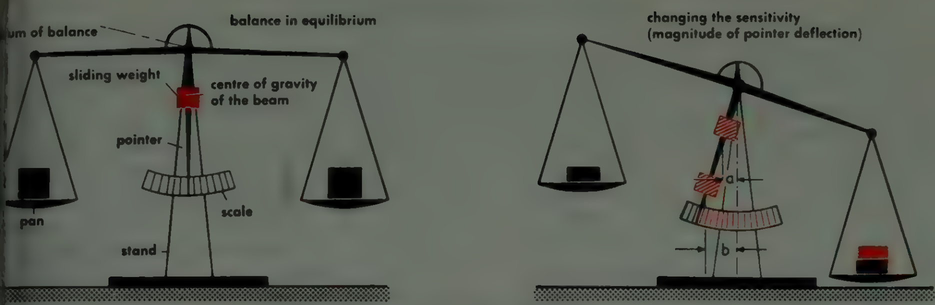


Fig. 1 EQUAL-ARMED BEAM SCALE

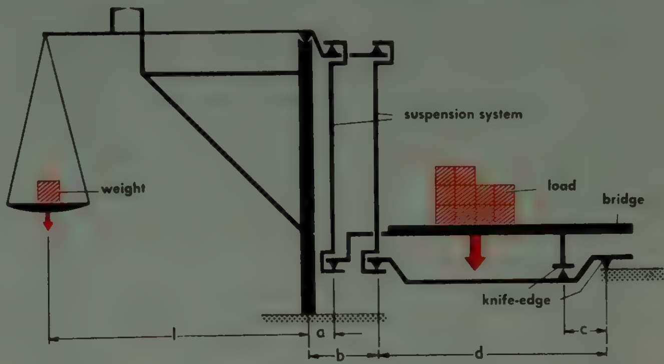


Fig. 2 WEIGHBRIDGE

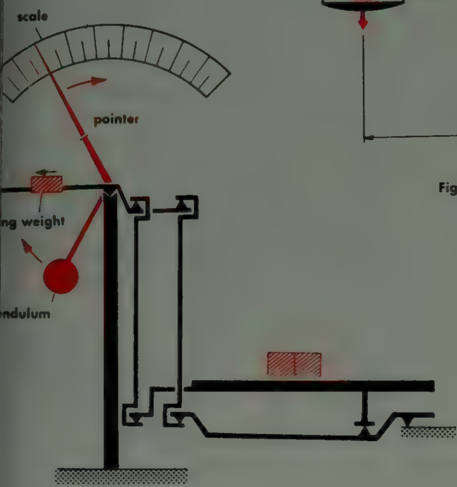


Fig. 3 PENDULUM TYPE WEIGHING MACHINE

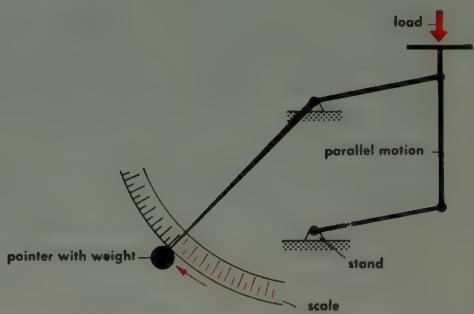


Fig. 4 LETTER BALANCE

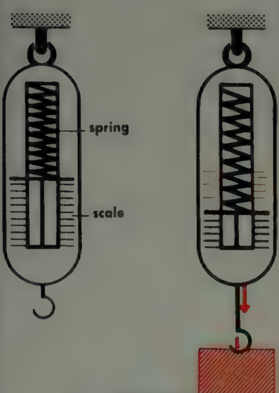


Fig. 5 SPRING BALANCE

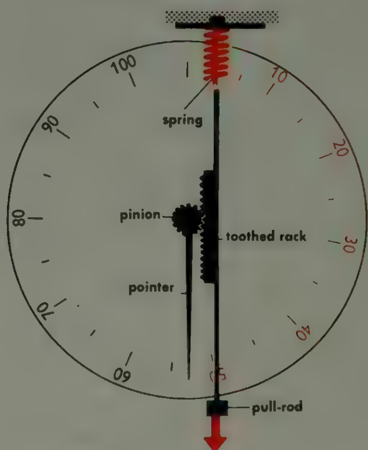
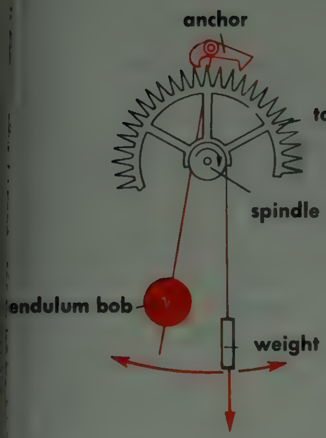


Fig. 6 SPRING BALANCE WITH CIRCULAR SCALE

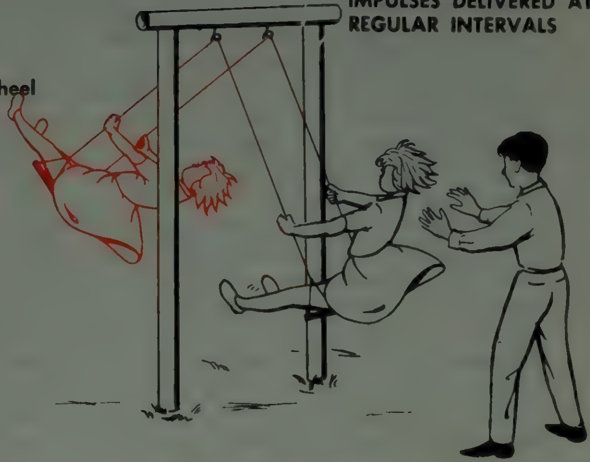
Any periodically repeated phenomenon can be utilised for time measurement, so long as the duration of the period remains accurately constant. In early timepieces the periodic movement was performed by a pendulum (Fig. 1). The weight which drives the watch is applied to the circumference of the spindle, causing it to rotate. This rotation is, however, arrested by the anchor, which is linked to the pendulum and which periodically engages with, and releases, a toothed wheel called the escape wheel (the combination of escape wheel and anchor is called the escapement). Each time the pendulum reaches its maximum amplitude, one of the projections (called pallets) of the anchor releases a tooth of the escape wheel, allowing this wheel to rotate a corresponding amount. Its rotation is therefore performed in a series of jerks, controlled by the anchor and pendulum, and this rotation is transmitted to the hands of the clock through a train of gear wheels. Friction would soon cause the pendulum to stop swinging if it were not given an impulse at regular intervals to keep it in motion, just as a child's swing has to be pushed each time it reaches it's full amplitude (Fig. 2). In the pendulum clock an impulse is imparted to the pendulum by the escape wheel (which is driven by the weight) through the pallets. The frequency (number of swings per second) of the pendulum can be varied by sliding the bob of the pendulum up or down on its rod. Lowering the bob makes the pendulum swing more slowly, and vice versa. In this way the period (time of oscillation) of the pendulum can be adjusted and, the clock itself thus be regulated. In watches the controlling action of the pendulum is performed by a device called the balance (Fig. 3). Attached to the spindle of the balance is a spiral spring, named the balance spring or hairspring, which controls the oscillations of the balance. Attached to the balance is a pin which engages with the lever. With each oscillation of the lever the pallets release the escape wheel, allowing it to rotate a distance corresponding to one tooth. At the same time an impulse from the escape wheel (which is driven by the mainspring) is transmitted to the balance through the lever and pin and thereby keeps the balance in motion. The function of the latter is thus entirely analogous to that of the pendulum in a pendulum clock. The type of escapement illustrated in Fig. 3 is the so-called lever escapement; it was invented about two hundred years ago, and is now widely employed; there are several other types of escapement for watches. The balance performs five to-and-fro movements per second, i.e., the second hand moves in five tiny jerks each second. The escape wheel drives the minute hand and hour hand through a train of gear wheels.

The transmission of the movement of the minute wheel and hour wheel to the hands of the watch is shown in Fig. 4. The minute wheel performs one complete revolution per hour, and so does the minute hand, which is mounted on the same spindle. During the same length of time the hour wheel and hour hand perform only one-twelfth of a revolution.



ig. 1 PENDULUM DRIVE

Fig. 2 PENDULUM MOTION IS MAINTAINED BY IMPULSES DELIVERED AT REGULAR INTERVALS



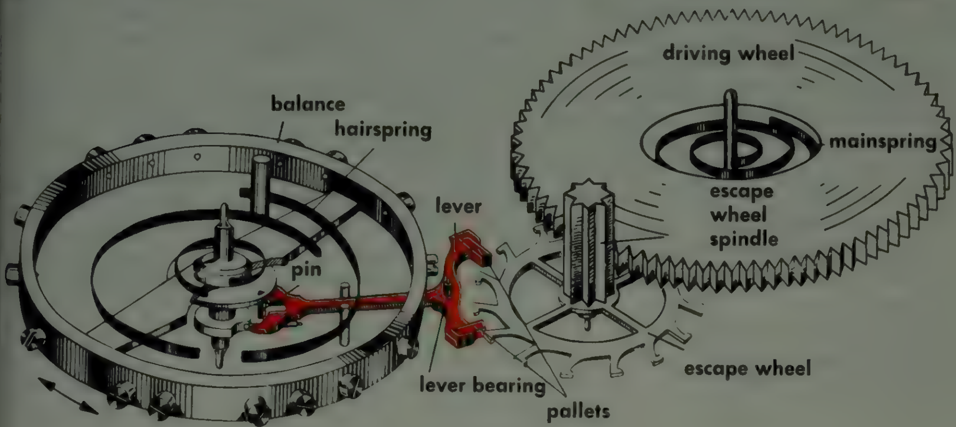


Fig. 3 DRIVE MECHANISM OF A WATCH

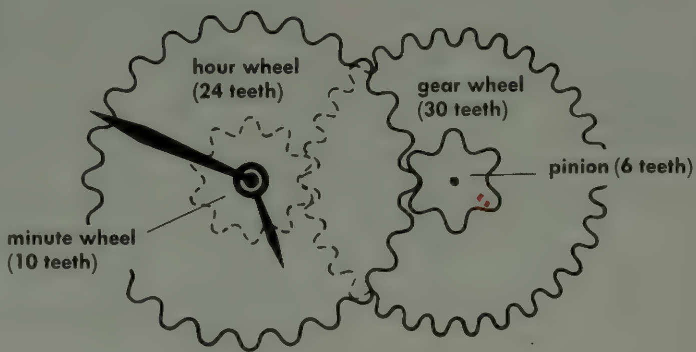


Fig. 4 TRANSMISSION OF MOTION FROM THE MINUTE AND HOUR WHEELS TO THE HANDS

In every method of measuring time a periodically recurring process is used as the basis of the measurement. In clocks and watches the periodic recurrence is provided by the swinging pendulum or the oscillating balance (see page 245).

Essentially, any periodic process can be used for controlling a timepiece. In the device called the "quartz clock" these consist of the "thickness vibrations" that quartz (and certain other crystals) perform under particular conditions. The principle of such vibrations can best be illustrated by a very simple example: when a jelly pudding is struck with a spoon (Fig. 1), it begins to wobble. These movements are a kind of thickness vibrations. If the jelly is struck at regular intervals corresponding to its vibration frequency, the amplitude of the vibrations will increase to such an extent that the pudding may actually break up. A quartz crystal cut in a certain way exhibits a similar effect. In this case, however, the excitation of the crystal is not done by mechanical impulses but by periodic electric charging (Fig. 2). For this purpose a phenomenon known as the *piezoelectric* effect is utilised: when the crystal is subjected to alternate compressive and tensile strains, opposite electric charges are produced on different faces; conversely, when electric charges are applied to these faces of the crystal, the latter undergoes expansion and contraction. By this means the crystal can be set vibrating. The frequency of these thickness vibrations depends solely on the dimensions of the crystal and can be given any desired value by appropriately choosing these dimensions. For a given set of conditions the frequency is extremely constant. A quartz crystal can thus be used as a highly accurate regulator for an electric oscillatory circuit (Fig. 3).

The quartz clock comprises a tube (or valve) transmitter (Fig. 4), or a transistorised transmitter (cf. page 88), whose oscillatory circuit is controlled by a quartz crystal. At the output end (A) an alternating voltage with a frequency possessing a high degree of constancy is obtained. This output can be fed through frequency reducing circuits, or be supplied to a high-frequency motor, and thus be used to drive a normal clock. The time-keeping accuracy is very high (to within one ten-thousandth of a second over a period of months). Quartz clocks have acquired importance as master clocks for public timekeeping purposes and for keeping the frequencies of radio transmitters constant.

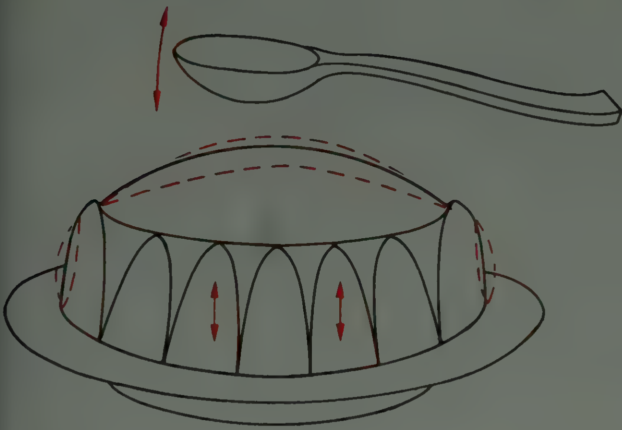


Fig. 1 *Thickness vibrations of an elastic substance*

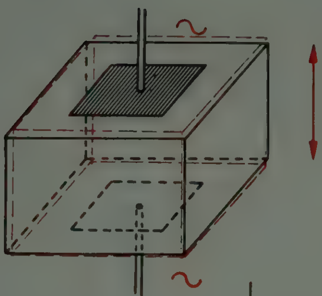


Fig. 2 *Thickness vibrations of a quartz crystal under the influence of an alternating voltage (oscillator quartz)*

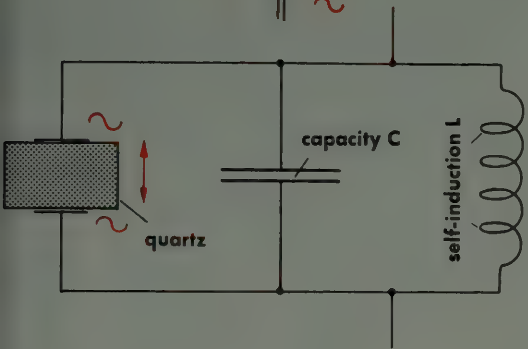


Fig. 3 *Control of an electric oscillatory circuit (L, C) by means of a piezoelectric crystal (oscillator quartz) so as to ensure constant frequency*

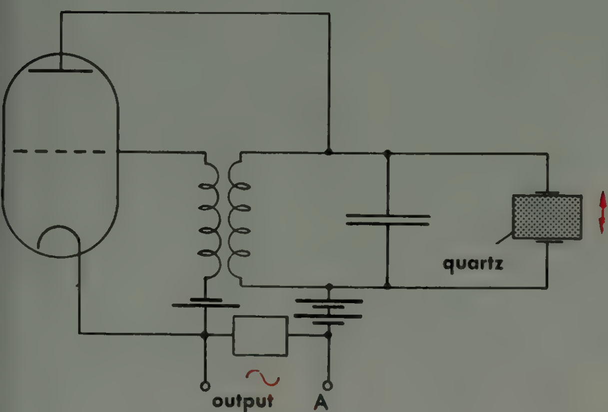


Fig. 4 *Principle of the quartz clock: the frequency of the electro-magnetic oscillations of a transmitter is kept constant by means of a piezoelectric crystal*

Miniature quartz clock by Patek Philippe.
Accuracy: 1/10 second per day
Photo Guy de Belleval

The atomic clock is the most accurate time measuring device at present known. Whereas the accuracy attainable with high-precision pendulum clocks is about 10^{-7} (i.e., an error of 3 seconds per year), the accuracy of the quartz clock (see page 216) is 100–1000 times greater. However, quartz clocks have the great disadvantage that their vibration frequency changes in course of time. Hence the frequency has to be checked from time to time and readjusted. This is done with the aid of an atomic clock, in which the frequency is determined by molecular vibrations and remains constant. The accuracy of the atomic clock is about ten times as great as that of the quartz clock.

The operating principle of an ammonia atomic clock is as follows:

Gaseous ammonia (NH_3) is introduced through the nozzle on the left in Fig. 1. The nozzle consists of a number of very fine parallel passages, and the molecules travelling to the right enter the focusing device, which consists of a number of metal cylinders (usually four) charged to a high voltage. Among the NH_3 molecules are some with a higher and some with a lower energy content. Both kinds of molecule have a dipole moment, i.e., they align themselves in an electric field in such a manner that the high-energy molecules dispose themselves against the direction of the field, whereas the low-energy molecules dispose themselves in that direction. Because of these properties of the NH_3 molecule, it is possible, with the aid of the very inhomogeneous field in the focusing device, to separate the high-energy molecules from the low-energy ones, so that the former are thrust towards the central area by the electric field, while the latter are thrust outwards (Fig. 2). The high-energy molecules are collected by the focusing device and directed into a cavity resonator (see page 86). The resonator is a metal box in which a stationary high-frequency wave can be formed with the aid of a feedback system. Activated by this wave, the high-energy NH_3 molecules acquire vibrations of about 24 milliard[1] cycles/sec. and give off their energy to the wave, which is thereby amplified. This high-frequency energy has a frequency which remains very accurately constant, thus providing the basis for the time measurement.

1. A milliard is equivalent to one billion in U.S.A.

Fig. 1 PRINCIPLE OF ATOMIC CLOCK
(viewed at right angles to molecule beam)

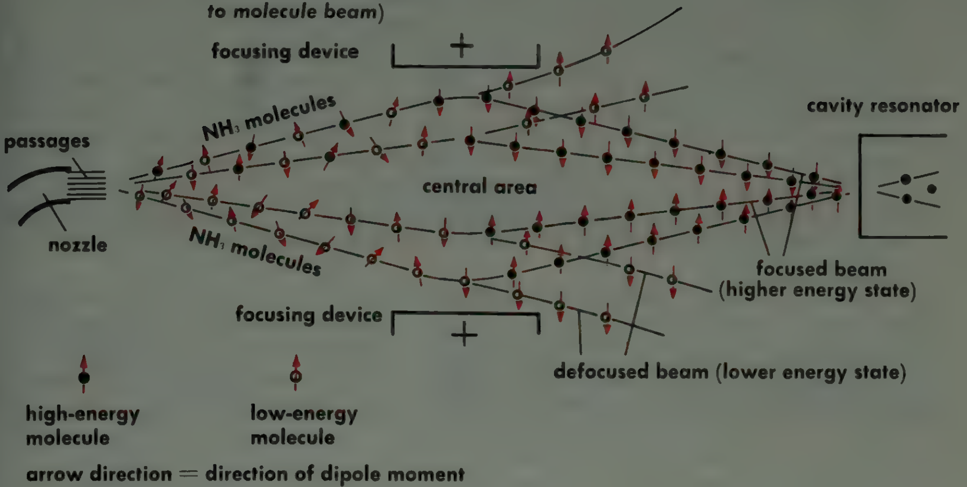
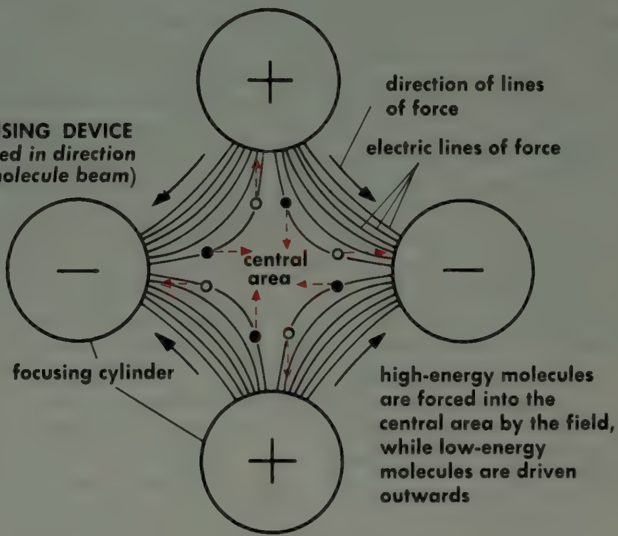


Fig. 2 FOCUSING DEVICE
(viewed in direction of molecule beam)



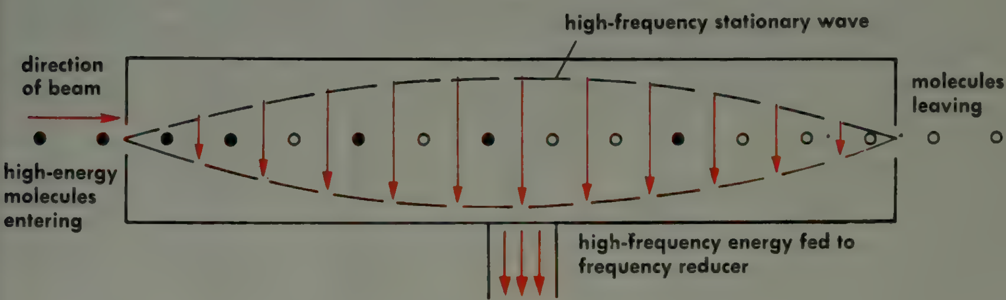


Fig. 3 CAVITY RESONATOR

A barometer is a device for measuring atmospheric pressure. The average atmospheric pressure at sea level is 1 atmosphere, which is the pressure that will support a column of mercury 760 mm (29.92 inches) high, i.e., about 1 kg/cm^2 (14.7 $lb./in.^2$). This would correspond to the pressure exerted by a column of air about 5 miles (8 km) high if its density were constant and equal to that at sea level. (Actually the depth of the atmosphere is substantially greater, as the density diminishes with increasing altitude). This explains why the atmospheric pressure on the top of a mountain is lower. If the mountain is, say, 10,000 ft. (3000 m) high, the column of air pressing down at the top is only about 3 miles (assuming constant density) (Fig. 1). Actually the atmospheric pressure at any particular altitude is never constant, but varies by relatively small amounts about the average. These variations provide data for weather forecasts.

When a rubber suction pad is pressed against a smooth wall (Fig. 2), it will remain adhering to the wall. By pressing the pad down flat against the wall surface the air is expelled from the cavity under the pad and a vacuum is formed there. The air pressure which originally acted in that cavity now no longer thrusts against the inside of the pad. When the latter is released, it will remain in the compressed position because the air pressure acts on the outside only and thus presses the edge of the pad down so firmly that no air will penetrate into the cavity, which therefore remains void of air. The external air pressure thus keeps the suction pad firmly pressed against the wall. Difference of air pressure is also utilised when we drink a liquid through a straw. Air is sucked out of the straw, and the atmospheric air pressure, which acts upon the surface of the liquid in the glass, will force the liquid up through the straw (Fig. 3). If a long glass tube, which is sealed at one end and open at the other, is filled with mercury and is then stood upright, with the open end downwards, in a dish containing mercury, then so much mercury will flow out of the tube until a column of mercury not more than 760 mm in height above the mercury surface in the dish remains (Fig. 4). The air pressure which acts upon the surface of the mercury in the dish is therefore able to hold up a 760 mm high mercury column. If the same experiment were carried out with water instead of mercury, the tube would have to be at least 33.9 ft. (10.33 m) high; and a column of air would have to be 5 miles high to exert a pressure of the same magnitude (Fig. 5). A mercury-filled tube over 760 mm in length as described above, represents the simplest form of the mercury barometer. The tube contains no air; the space above the mercury column is a vacuum.

Modern barometers are mostly of the aneroid type. An aneroid barometer (Fig. 6) consists of a hermetically sealed metal box, exhausted of air. The top and bottom of this box are thin corrugated plates. In the interior of the box is a spring which strives to force the top and bottom plates apart, against the external air pressure. If this air pressure decreases, the springs can expand; if it increases, the spring is compressed. The movement of the top plate is transmitted through a train of levers to a pointer which indicates the atmospheric pressure on an appropriately graduated scale.

The unit of atmospheric pressure usually employed in meteorology is the millibar. It is approximately equal to $\frac{1}{32}$ inch of mercury (actually a mercury column of 750 mm corresponds to 1000 millibars).

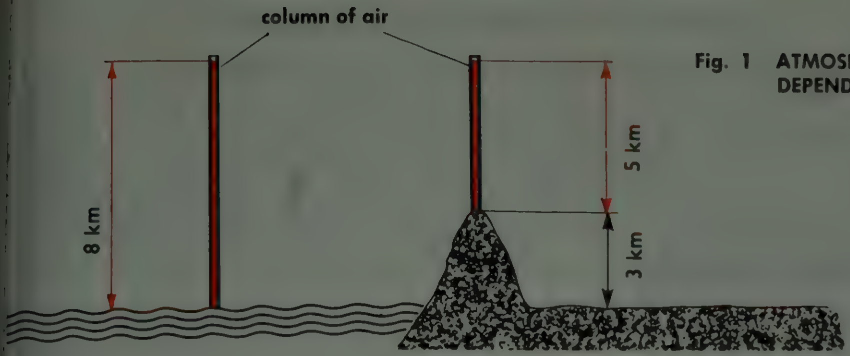


Fig. 1 ATMOSPHERIC PRESSURE DEPENDS ON ALTITUDE

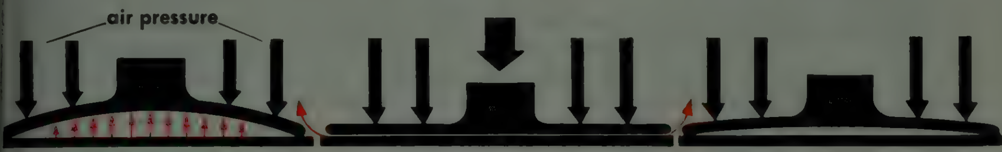


Fig. 2 RUBBER SUCTION PAD

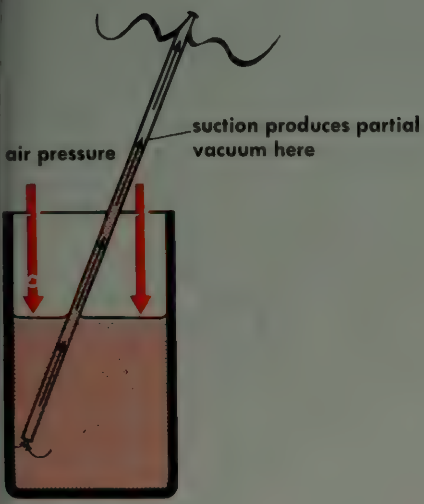


Fig. 3

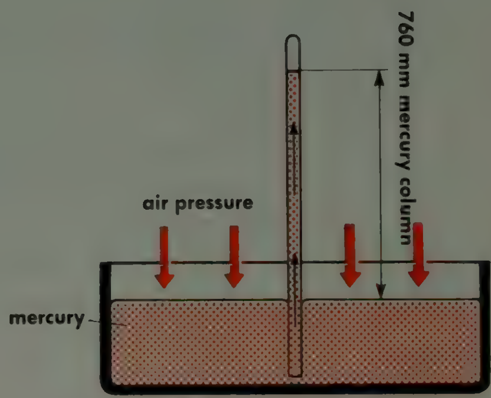


Fig. 4 PRINCIPLE OF THE MERCURY BAROMETER

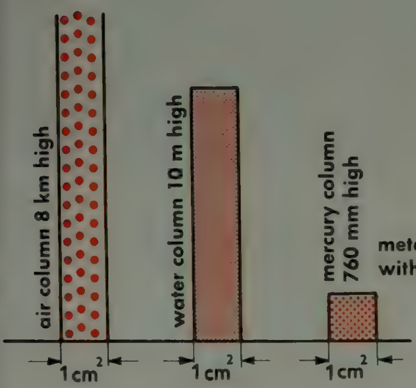


Fig. 5

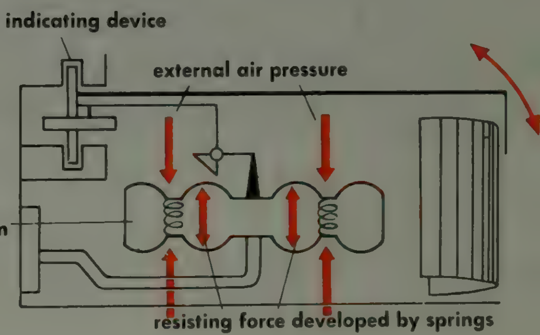


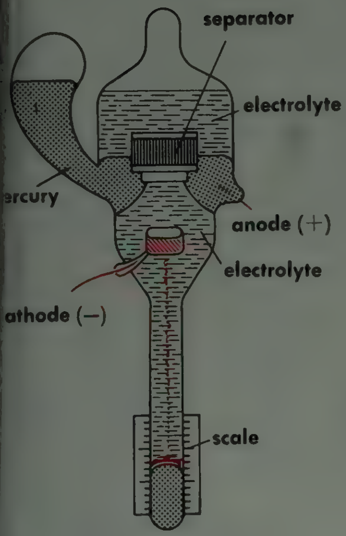
Fig. 6 ANEROID BAROMETER

Electricity meters are used for measuring the "quantity" of electricity consumed in houses and factories. What the meter actually measures is the electric energy that passes through the circuit. For direct current the energy is $A = V.I.t$, and for single-phase alternating current it is $A = V.I.t \cos \phi$ (cf. page 82), where V is the voltage, I is the current strength, t is the time, and $\cos \phi$ is the power factor. In the case of direct current the energy can be metered by determining the value of $I.t$ by means of an ampere-hour meter, so long as the voltage V in the mains is kept constant. In alternating-current supply systems, however, only watt-hour meters are used.

Fig. 1 illustrates the principle of the electrolytic meter, which can be used only for alternating current. It is now obsolete, but its mode of functioning is of some interest. A closed glass vessel contains an electrolyte consisting of a solution of a mercury salt. Part of the consumer current flows through this electrolyte; at the cathode a quantity of mercury proportional to the value of $I.t$ is deposited. The mercury collects in the bottom part of the vessel. The scale is graduated in kWh. The mercury can subsequently be retrieved and returned to the top compartment.

Fig. 2a shows a motor-type ampere-hour meter. It is in effect a small direct-current motor. The armature, or rotor, is an aluminium disc mounted in the field of a permanent magnet. The disc contains three coils which are supplied with current from a commutator consisting of three sectors. The current to the rotor coils is proportional to the total current passing through the consumer installation, from which it is branched off (Fig. 2b). Since the field strength of the permanent magnet is constant, the speed of rotation of the rotor is proportional to the current strength. The counting mechanism connected to the rotor shaft counts the revolutions, which correspond to the product $I.t$. If, instead of the permanent magnet, an electromagnet is used to produce the magnetic field, and if the consumer voltage V is applied to this coil, then the speed of rotation of the rotor is proportional both to V and to I, and therefore to the product $V.I$. In that case the counting mechanism directly records the energy $V.I.t$. The ampere-hour meter has thus become a true watt-hour meter.

The kind of meter nowadays generally used for alternating current is the induction meter (Fig. 3). It has no commutator. There are two electromagnets. The coil of one of these is energised by the consumer current; the other magnet coil is connected to the consumer voltage. If the current and the voltage in the consumer circuit are in phase with each other (see page 82), then the current in the voltage coil, and therefore the magnetic field of this coil, will have a lag of one-quarter period (90°) in relation to that of the current coil. The interaction of the two coils produces a moving magnetic field which induces eddy currents in the light aluminium rotor disc. These currents cause the disc to rotate in the direction of motion of the moving field. The speed of rotation of the disc is proportional to the strengths of the two magnetic fields, but it is also dependent upon the phase displacement of these fields (and therefore upon the power factor $\cos \phi$). This will readily be understood when one considers the case where the consumer installation has a 90° phase displacement between V and I (i.e., $\cos \phi = 0$); when this happens the two magnetic fields in the meter will be in phase, so that the rotor will then cease to rotate. The braking magnet (on the right in Fig. 3) constantly produces eddy currents in the rotor; these damp the rotation and thereby ensure that the rotor stops instantly when the consumer current ceases to flow.



ig. 1 ELECTROLYTIC METER

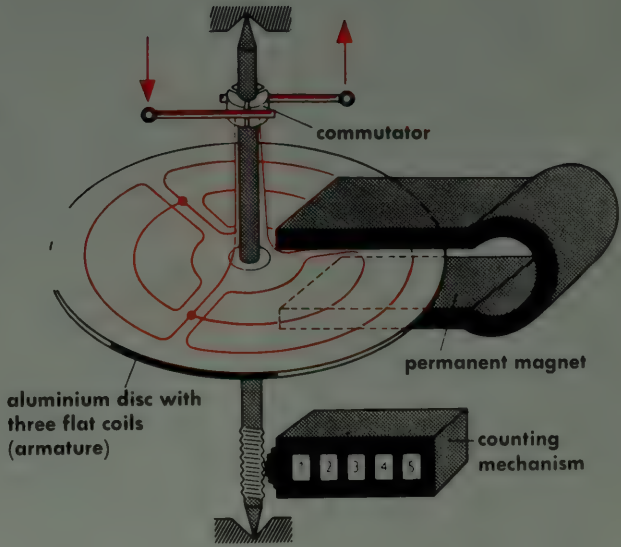


Fig. 2a MOTOR-TYPE METER (*schematic*)

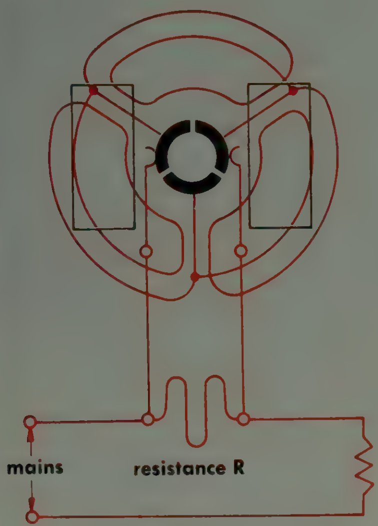


Fig. 2b WIRING DIAGRAM OF A MOTOR-TYPE METER

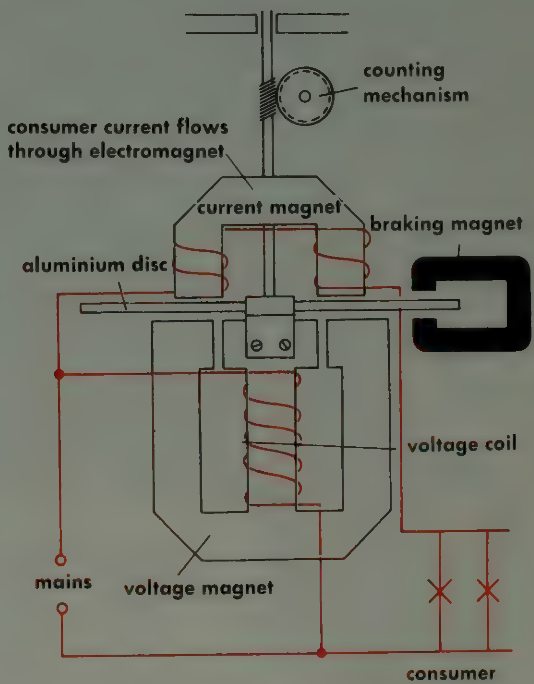


Fig. 3 INDUCTION METER

Electricity meter, used in France
Photo Sodel-Electricité de France


3
C C E - COMPTEURS "BT"
Fabr a Issy-Les Moulineaux.Seine
kWh
TYPE B1C3
MONOPHASE 2 FILS
10-30 A 220 V
C = 3,6 Wh/tr
N°

Gas meters are either of the "dry" or the "wet" type. The "*dry*" *gas meter* comprises two chambers of known volume, which are separated from each other by a leather diaphragm. Both chambers have inlet and outlet valves. The pressure of the gas causes the chambers to be alternately filled and emptied, whereby the diaphragm is alternately stretched and relaxed. These movements of the diaphragm are transmitted to a lever system which controls the valves of the chambers. The filled chamber controls the mechanism of the other chamber, which is about to be filled.

In order to keep the flow of gas through the meter as continuous as possible, two two-chamber systems are combined into a unit. Figs. 1a–c illustrate some stages in the alternate filling and emptying operations of the system. In Fig. 1a, measuring chamber Ia is full of gas. Just previously to this the control mechanism of chamber I effected the change-over of the valve so as to let gas flow into chamber IIb. When this chamber has filled up, the valve *a* opens chamber Ib further. It closes the latter again as soon as chamber IIa begins to fill up (Fig. 1b). Fig. 1c shows chamber Ib completely closed by the valve, while chamber IIa is in the final stage of filling. In this stage chamber Ib begins to empty its contents; at the same time, the diaphragm, acting through the crank mechanism, has set the valve *a* so as to let gas into chamber Ia. As chamber Ia fills up, the diaphragm *a* so actuated the valve *b* that chamber IIa is emptied and chamber IIb fills up.

The rate of emptying and filling of the chambers corresponds to the rate of rise and fall of the diaphragms; this rate, which is therefore a measure of the volume of gas consumed, is transmitted to a counting mechanism. This type of meter is able to measure even very small rates of gas flow with considerable accuracy, but at high rates of flow the accuracy diminishes because the valves can then no longer be changed over quickly enough. The meter is very reliable and unlikely to develop disorders, and for this reason it is widely used for metering domestic supplies.

The "*wet*" *gas meter* operates on the same principle as the drum type rotary meter for liquids (see page 260). The Crossley meter is an extensively used type. Its principle is illustrated in Fig. 2. The drum, which contains four compartments with inlet and outlet slots or holes, rotates half immersed in water, which acts as a seal. As soon as the inlet hole of a compartment emerges from the water, gas flows into it and fills the compartment as the water level in it goes down. When the outlet hole of the compartment emerges from the water, the gas flows out through this hole. Residual gas remaining in the compartment is forced out by the inflowing water as the drum continues its rotation.

In Fig. 2, compartment I of the meter is completely filled with water. The gas is being forced out of compartment II by the water. Compartment III is full of gas, and compartment IV, whose inlet hole has just emerged from the water, is beginning to fill up with gas. The gas pressure causes the rotation of the drum. The number of revolutions is recorded by a counting mechanism whose readings show the amount of gas consumed.

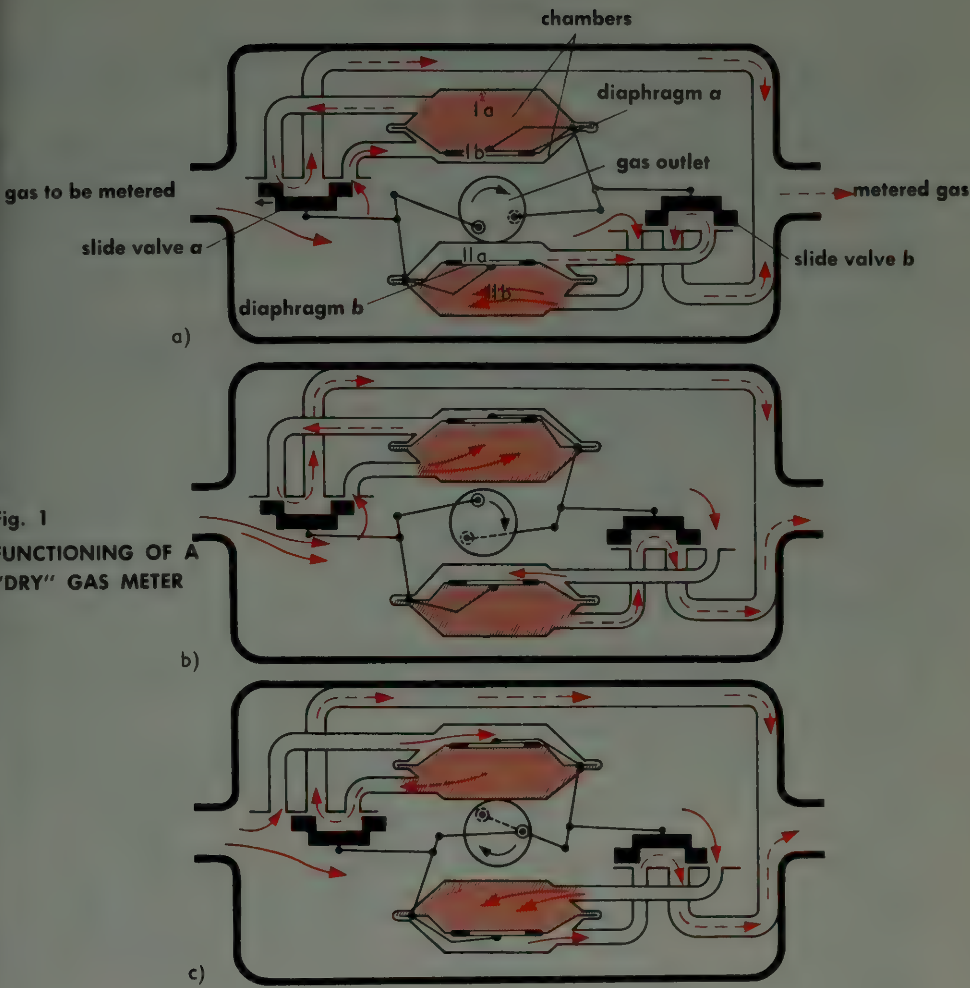


Fig. 1
FUNCTIONING OF A "DRY" GAS METER

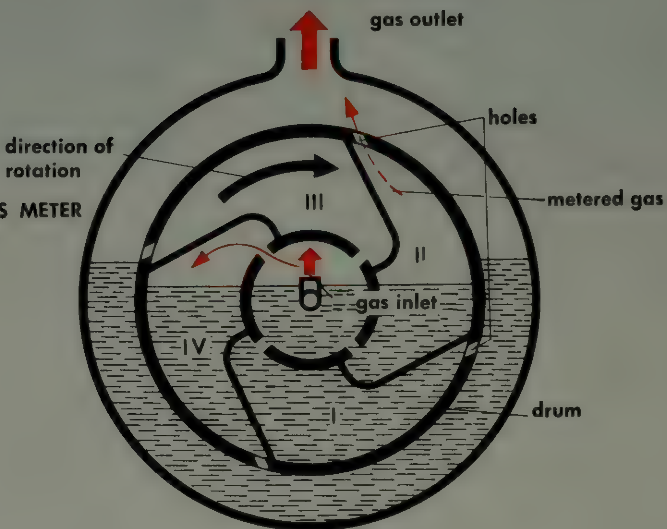


Fig. 2 "WET" GAS METER

Water meters used for measuring water consumption may function according to either of two different principles:

Rate-of-flow meters:

In devices of this kind the flowing water rotates a propeller whose speed of rotation (r.p.m.) is measured. A certain flow rate (cubic feet per second) corresponds to a certain speed of the propeller. The propeller speed indicator can therefore be calibrated directly in cubic feet per second (or gallons per second). A counting mechanism which totals the number of propeller revolutions will thus show the quantity of water consumed.

Volumetric meters:

In devices of this category a rotating chamber or container of known volumetric capacity is constantly filled with water and then empties itself. Here again the number of rotations is a measure of the quantity of water that flows through the meter. Volumetric meters are more particularly suitable for measuring low rates of flow. In the drum type meter (Fig. 1) the water enters the drum at the central inlet and fills one of the measuring compartments. The compartment is so designed that, as it fills up, the weight of the water makes the drum rotate (in Fig. 1 the left-hand deeper part of the compartment, when full of water, is heavier). The rotation of the drum causes the the water to flow out of the outlet opening. Meanwhile the next compartment is already being filled with water entering at the central inlet.

For the metering of domestic water supplies the rotating impeller type of meter is frequently employed. These devices are of two types. In one type (Fig. 2) the impeller is immersed in the flow of water, whereas the counting mechanism and dial are housed in a dry compartment. In the other type (Fig. 3) the counting mechanism and dial are also immersed. The latter type has better accuracy, but has the disadvantage that dirt in the water is liable to clog up the dial.

Woltmann meters are used for the metering of larger quantities of water, especially in industry (Fig. 4). This kind of meter consists essentially of a straight length of pipe in which a propeller is installed, which rotates and transmits its motion through a worm gear.

A different type of metering device is the Venturi meter (Fig. 5). It is based on the principle that different flow velocities produce different amounts of suction. The Venturi meter comprises a pipe formed with a constriction or "waist". The flow velocity at B is higher than at either of the sections A, and the suction (measured by the difference in level in the liquid in the two legs of the U-tube manometer) at B is correspondingly greater. Since the difference in pressure between B and A depends on the flow velocity, it must also depend on the quantity of water passing through the pipe per unit of time (flow rate in cu. ft./sec. = cross-sectional area of pipe in ft.^{2} × flow velocity in ft./sec.). Hence this pressure difference provides a measure for the flow rate. In the gradually tapered portion of the pipe downstream of B the velocity of the water is reduced and the pressure in the pipe restored to the value it had before passing through the constriction.

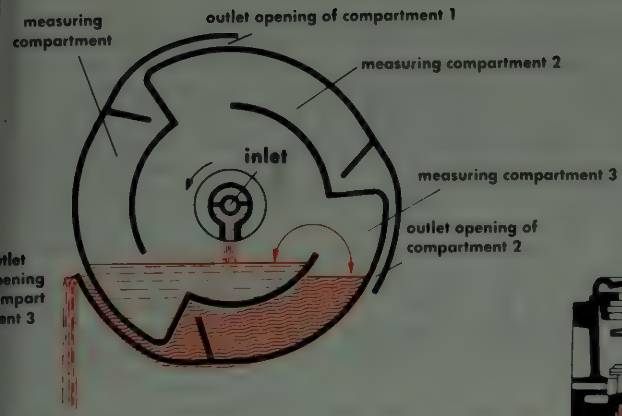


Fig. 1 DRUM TYPE METER (*volumetric meter*)

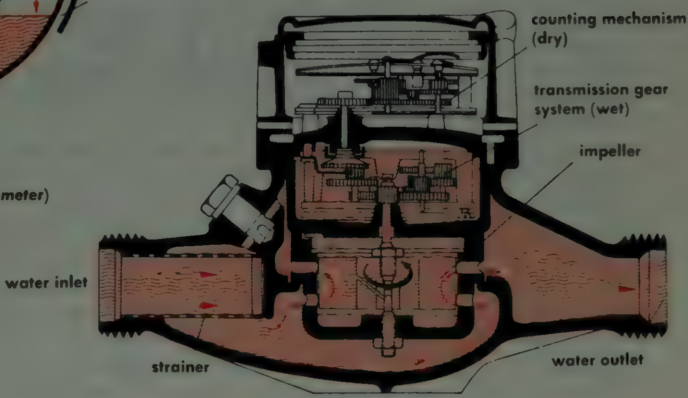


Fig. 2 IMPELLER TYPE METER WITH "DRY" COUNTING MECHANISM (*rate-of-flow meter*)

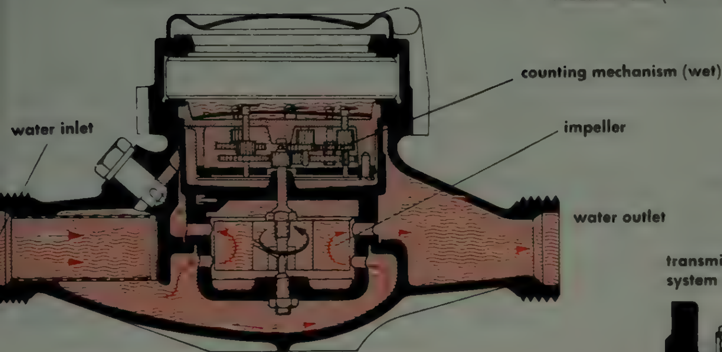


Fig. 3 IMPELLER TYPE METER WITH COUNTING MECHANISM IMMERSED (*rate-of-flow meter*)

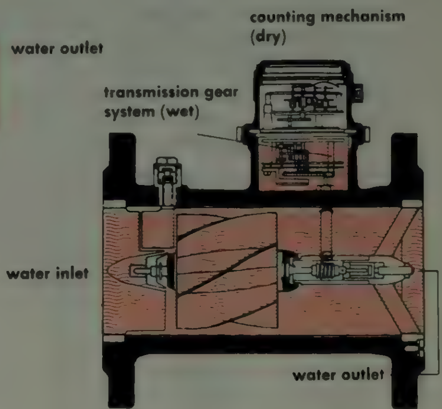


Fig. 4 WOLTMANN METER (*rate-of-flow meter*)

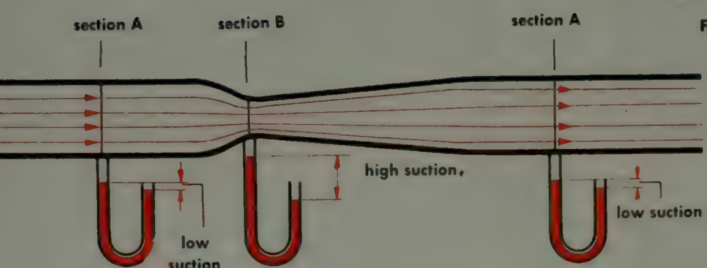


Fig. 5 VENTURI METER

If two bodies with different temperatures can exchange heat with each other, they will finally both acquire the same temperature at some intermediate value between their original temperatures. There are three kinds of heat transfer: convection, radiation, and conduction. In the case of convection (Fig. 1), heat transfer is effected by the flow of a gas or a liquid caused by local heating and the effect of gravity (e.g., when a room is heated by warm air rising from a stove or a radiator, water heating; cf. page 274 *et seq.*). Thermal radiation is independent of heat-transmitting matter and can also be effected through vacuum. Every hot body emits rays (electromagnetic waves) (Fig. 2) which are located in the infra-red portion of the spectrum, i.e., they have wavelengths longer than those of visible light rays. Bodies which absorb a large amount of radiation (e.g., soot) also re-emit a large amount of radiation. On the other hand, a suitably transparent body will, in general, allow radiant heat to pass and will undergo only a slight rise in temperature in consequence.

Transfer of heat by conduction occurs in the interior of solid bodies (Fig. 3) and also in liquids and gases. The rate of transfer will depend on the thermal conductivity, the thermal capacity of the materials, and the temperature difference. In building construction the thermal conductivity is a criterion for the heat insulating capacity of a material, which is important with regard to the loss of heat through the walls of a building. In metric units, thermal conductivity is expressed as the amount of heat (kcal)[1] which flows through 1 m^3 cube of a material in a time of 1 hour when the temperature difference between the two opposite faces of the cube is constantly maintained at 1° C. The "cube" must be conceived as a 1 m^2 area of a 1 m thick wall through which the flow of heat occurs. Steel has a thermal conductivity of about 50 kcal/m^3h °C; the corresponding figure for natural stone is about 2.5, for concrete about 1.5, for brickwork about 0.75, for glass about 0.70, for wood about 0.15, and for cork panels about 0.035. In British units the conductivity is usually expressed as the heat (in B.T.U.[2]) that is transmitted through 1 sq.ft. of a 1 ft. thick wall in 1 hour for a temperature difference of 1 F between the two wall faces.

The heat capacity (kcal per °C) of a body is the amount of heat that must be added to it, or abstracted from it, so as to produce a change of 1° C in its temperature. The thermal resistivity is the reciprocal of the thermal conductivity and provides a direct indication of the heat-insulating power of a material: the thicker the layer of material and the lower its thermal conductivity, the higher will be its resistivity (Fig. 4).

The best heat insulator is a vacuum (in combination with efficient protection against heat transfer by radiation). An important application is, for example, the vacuum flask, a double-walled glass vessel, the cavity between the two walls being exhausted of air (and silvered to minimise losses of heat due to radiation), so that this vacuum serves to insulate the contents of the flask and keep them hot or cold. In building construction, heat insulation is obtained by using materials which are poor conductors of heat, especially materials containing a large number of air-filled voids or pores. Since air itself is a poor conductor of heat, such materials are good insulators. Air as a heat-insulating substance is also utilised in cavity walls, i.e., brick walls consisting of an outer and an inner "leaf" separated by an air gap.

1. Kilo calories.
2. 1 B.T.U. = 0.252 kcal.

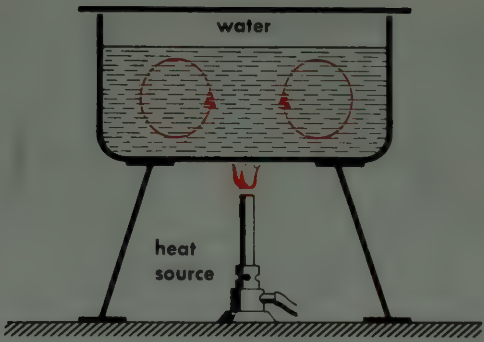


Fig. 1 CONVECTION

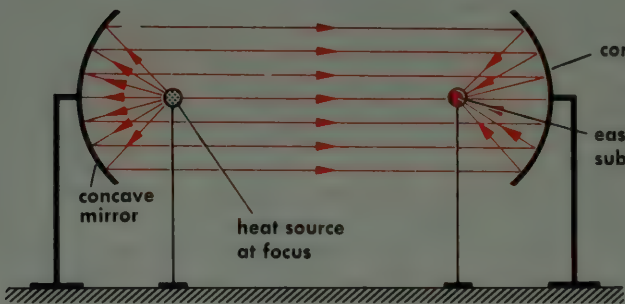


Fig. 2 RADIATION

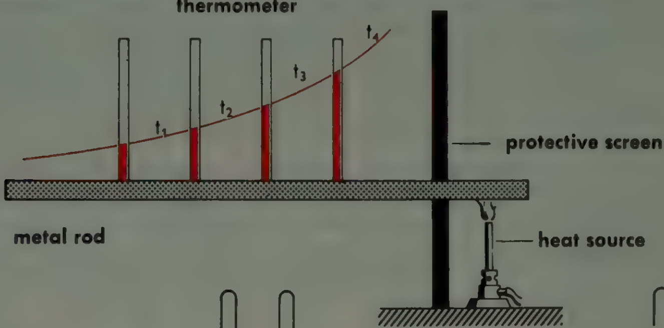


Fig. 3 CONDUCTION

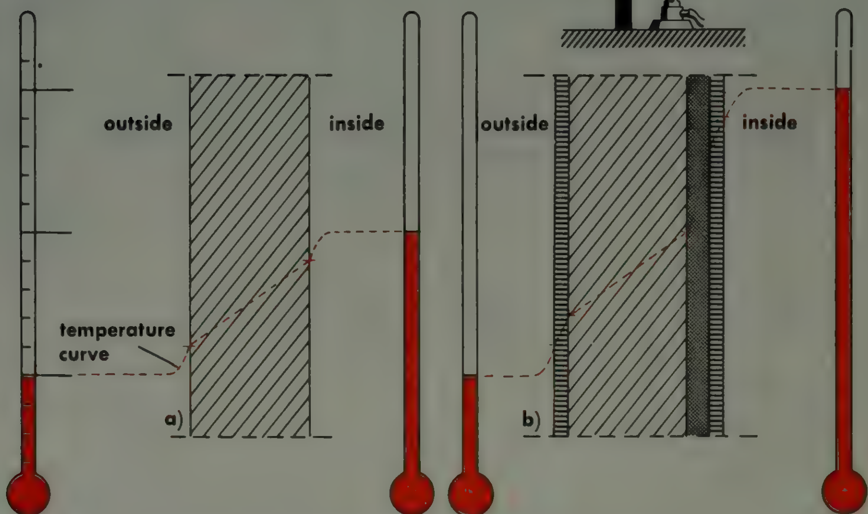


Fig. 4 EXAMPLES OF TEMPERATURE GRADIENTS IN A WALL:
(a) without insulation; (b) with insulating backing and facing

A block and tackle arrangement is an extensively applied mechanical device which is used for lifting loads and hauling. It comprises two or more pulley blocks and a length of rope that passes round the pulleys. With its aid it is possible to lift heavy loads by the application of only a relatively small force. The mechanical principle of equilibrium based on the equation: force × arm of the force = load × arm of the load is illustrated in Figs. 1a and 1b. In Fig. 1a there is obviously no equilibrium, but in Fig. 1b the equilibrium equation is satisfied, i.e., one boy at a distance from the pivot of the see-saw (the fulcrum of the pivot) will counterbalance two boys (of the same weight as the other) at a distance $\frac{1}{2}\,l$ from the pivot. It will also be apparent that if the boy on the right-hand side descends a distance s, the boys on the left will rise a distance $\frac{1}{2}\,s$. The product of the force (i.e., the weight moved) times the distance is therefore the same on both sides. This product—force × distance—is called "work". A similar situation presents itself when a rope carrying a weight at each end is passed round pulleys. In Fig. 2 there is no equilibrium, since the left-hand load is twice as large as the right-hand load and the lever arm (equal to the radius of the pulley) is the same for both loads (just as in Fig. 1a). In the arrangement shown in Fig. 3 there is a top "fixed" pulley and a moving pulley, the latter being suspended by two "falls" of rope, in each of which acts a pull equal to half the load. The sole function of the top pulley is to change the direction of the rope so that the force needed on the right to balance the left-hand load can act as a downward pull. This arrangement therefore enables a load of a certain weight to be raised by the application of a force only half as great as that load. Thus a mechanical advantage is gained, but there is no saving in the amount of work, because in order to raise the load a certain distance, the right-hand end of the rope (where the lifting force is applied) must move twice that distance. The product of force and distance i.e., the amount of work done, is therefore the same in both cases.

By increasing the number of pulleys and falls of rope, the mechanical advantage is increased according to $W = nP$, where W is the load, P is the force applied, and n is the number of falls of rope that support the moving block. Thus, in the arrangement illustrated in Fig. 6 there are six falls, enabling a load to be lifted by the application of a force only one-sixth as large.

Different arrangements are possible. In the type of system illustrated in Fig. 4 (which can be conceived as derived from Fig. 3 as the fundamental case) each additional moving block doubles the mechanical advantage, i.e., $W = 2^m P$, where m is the number of moving blocks. Thus, in Fig. 4 we have $W = 2^2 P = 4P$ or $P = \frac{1}{4}\,W$. If there were a third moving pulley similarly suspended, the force would be one-eighth of the load, and if there were four moving pulleys, it would be one-sixteenth, etc.

The arrangement illustrated in Fig. 5 is known as a differential pulley block. The top block comprises two firmly interconnected pulleys of different diameter. The rope passes continuously round the two pulleys shown. This arrangement provides a mechanical advantage determined by the difference in diameter $(R - r)$ between the two pulleys.

Fig. 1a NO EQUILIBRIUM

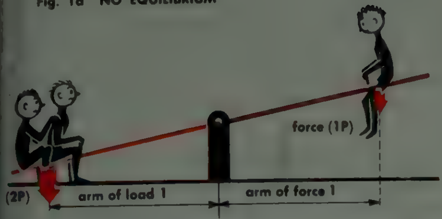
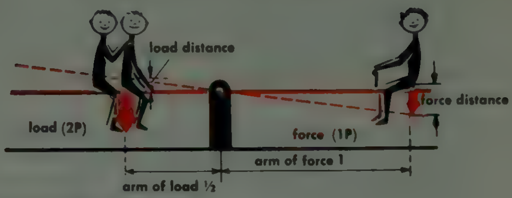


Fig. 1b EQUILIBRIUM
(force x arm of force load x arm of load)



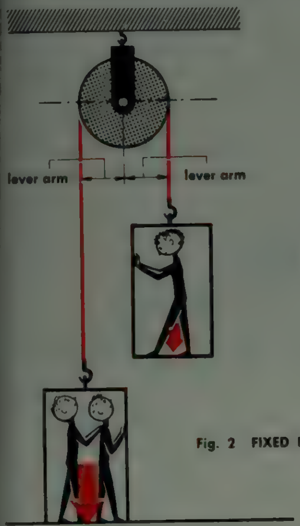
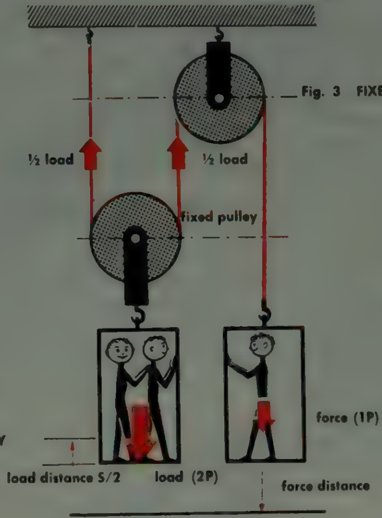


Fig. 2 FIXED PULLEY

Fig. 3 FIXED PULLEY AND MOVING PULLEY



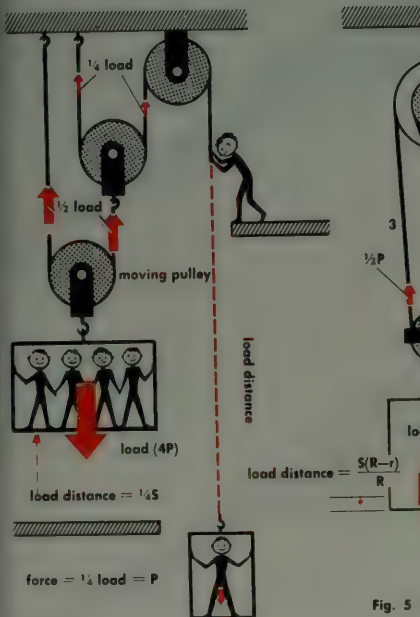


Fig. 4 RATIO OF FORCE TO LOAD AND OF THE DISTANCES THROUGH WHICH FORCE AND LOAD ARE MOVED

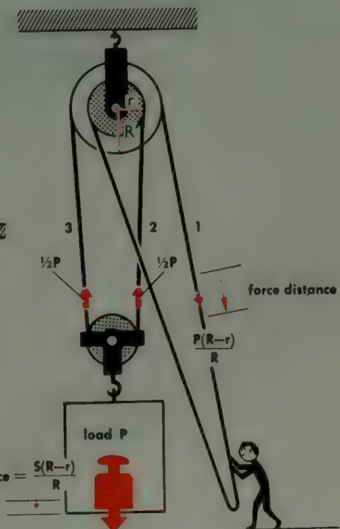


Fig. 5 DIFFERENTIAL PULLEY BLOCK

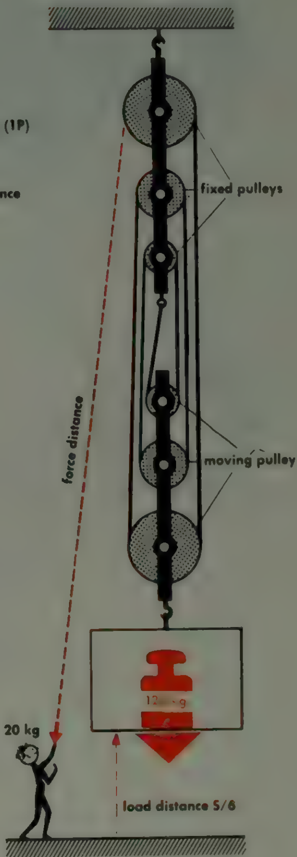
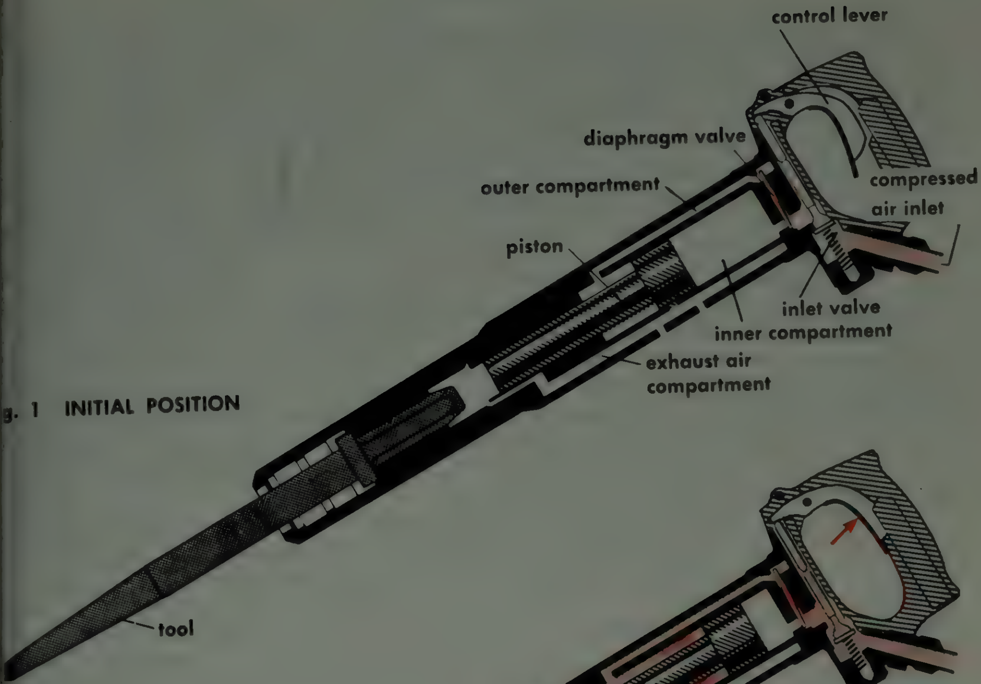


Fig. 6 MULTIPLE BLOCK AND TACKLE

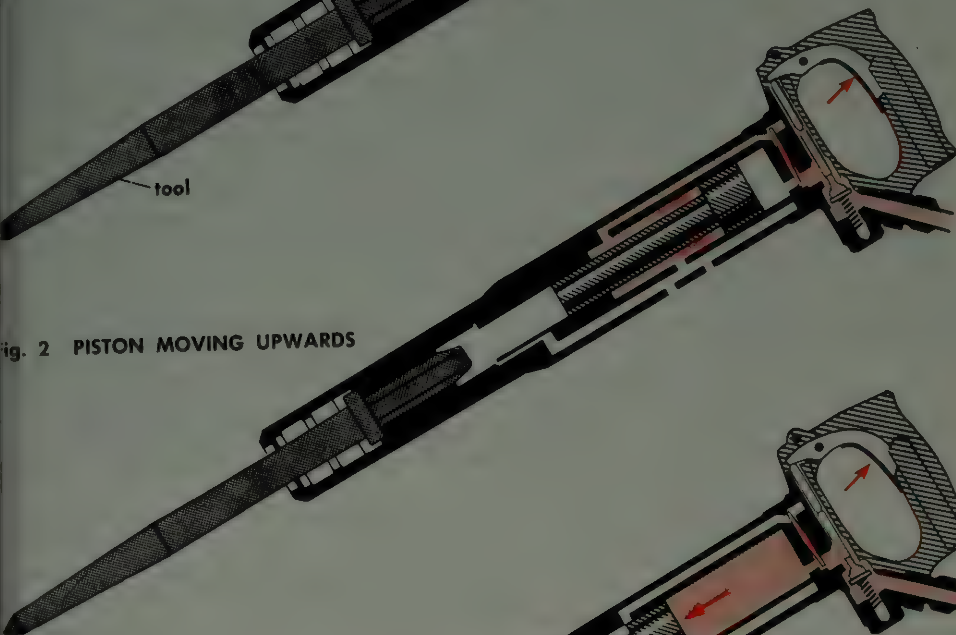
A pneumatic hammer delivers a large number of blows in rapid succession These blows are delivered by a piston which moves to and fro in a cylinder and is worked by compressed air. The piston strikes the upper end of a tool inserted in the front of the appliance. Various kinds of tool can be fitted, so that the appliance can be used not merely as a hammer, but as a pick, a concrete breaker, a digging tool, etc.

The compressed air is supplied through a hose from a compressor (see page 32) When the inlet valve is opened by means of the lever mounted within the handle (Fig. 2), the air flows through the diaphragm valve into the outer air compartment The diaphragm valve is convex, so that it can rock to and fro. In Fig. 2 this valve is in the position where it opens the inlet passage to the outer compartment. From here the air flows into the inner compartment from below and forces the piston upwards. The air above in the space above the piston undergoes compression, so that it forms an air cushion which somewhat softens the impact of the return movement of the piston. At the same time, this air acts against the underside of the diaphragm valve and tilts it the other way, with the result that the inlet passage to the outer compartment is closed and the inlet passage to the inner compartment is opened (Fig. 3). The compressed air is thereby admitted to the top of the piston and forces it downwards, so that it strikes the upper end of the tool. Both during the downward and during the upward stroke of the piston the expanded exhaust air is discharged into the exhaust air compartment and thence into the open air.

All pneumatic tools of the hammer or drill type function on the principle described here, though there are differences of detail in the design of the valves.



g. 1 INITIAL POSITION



ig. 2 PISTON MOVING UPWARDS

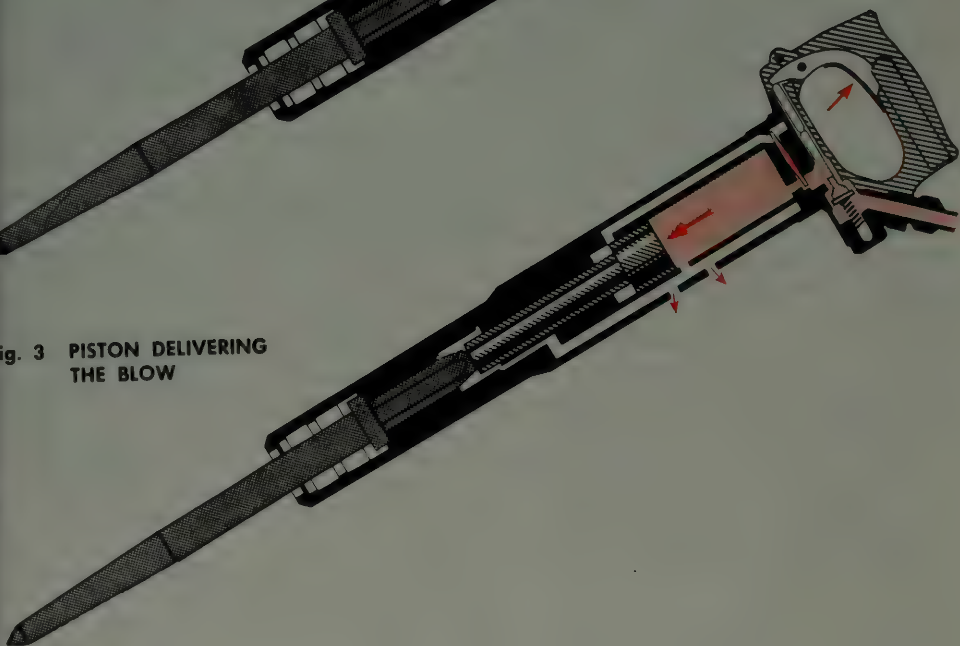


Fig. 3 PISTON DELIVERING
THE BLOW

Cranes are of many different kinds, depending on the purpose for which they are intended and on the magnitude of the loads to be handled. They may be mobile or stationary. The load is picked up by means of such attachments as hooks or tongs (for individual loads) or by means of buckets, skips or grabs (for bulk materials). A special lifting attachment, used mainly for the handling of scrap iron, is the lifting magnet. These attachments are usually suspended from wire ropes which pass round various pulley systems and are wound on hoist drums which are driven by electric motors. The principles of the pulley block (see page 264) are often embodied in the lifting tackle of cranes.

A major distinction can be made between bridge cranes and jib cranes. A crane of the bridge type has a trolley, or crab, which travels along a track and carries the winch that lifts the load. The commonest form of bridge crane is the overhead travelling crane used in factories, workshops, etc. Such a crane usually consists of a "bridge", comprising two girders, each end of which is mounted on a truck which travels on an overhead track extending the length of the building. The crab can travel to and fro on a transverse track installed on these girders. The combined movements of the crab and of the crane as a whole enable the lifting hook to be brought into position at any desired point in the building.

A jib crane has an arm — called the jib (or boom) — which can usually perform a "slewing" motion, i.e., rotate horizontally about a vertical pivot (the "king pin") mounted in the substructure of the crane. In addition, the jib is often able to perform a "luffing" (or "derricking") motion, i.e., it can be raised or lowered by varying its angle of inclination. The combination of these two motions enables the hook to be brought into position at any desired point within a certain radius. A type of jib crane used on construction sites for tall buildings is the tower crane (Fig. 1). Another type of jib crane is the so-called level-luffing crane, which is so designed that, by means of some form of compensating mechanism, the load moves in a horizontal path when the jib is luffed (raised or lowered); this arrangement has certain technical and operational advantages, especially in quay cranes used for loading and unloading ships at ports. Fig. 2 shows a particular form of level-luffing crane which has what is known as a double-lever jib. In this crane the compensating action is obtained by the movements of the jib lever. An advantage of this crane is that its projecting jib lever provides a greater amount of lateral clearance than an ordinary straight jib, so that there is ample space for handling bulky loads.

Every crane has a certain lifting capacity, ranging from a few tons to many hundreds of tons, depending on the type of crane and the purpose for which it is intended. In jib cranes the capacity usually varies with the radius, which depends on the slope of the jib. When the latter is raised to a steep slope, the radius — the distance from the load to the centre of the king pin — is small, and the crane can then carry a heavier load than when the jib is lowered to its farthest extent and the radius is large. This difference in lifting capacity at different radii is determined by the stability of the crane, i.e., its safety against overturning. The weight of the load (suspended from the jib) multiplied by the radius constitutes the overturning moment. The latter is counterbalanced by a heavy counterweight which is located a certain distance rearward from the king pin and develops a counterbalancing moment. This counterweight may be mounted on the substructure or on a special secondary jib projecting to the rear and is sometimes movable, so that the counterbalancing moment can be varied within certain limits. The overturning moment must always be smaller than the counterbalancing moment, and for this reason only a certain maximum load is permissible at a certain radius.

Besides hooks, a variety of lifting and handling devices can be attached to cranes. An important device for picking up bulk materials such as coal, ore, etc. is the grab (Fig. 3). It consists of two shells which can open and close to pick up the load and subsequently discharge it. These movements are produced by the actuation of the holding rope and the closing rope.

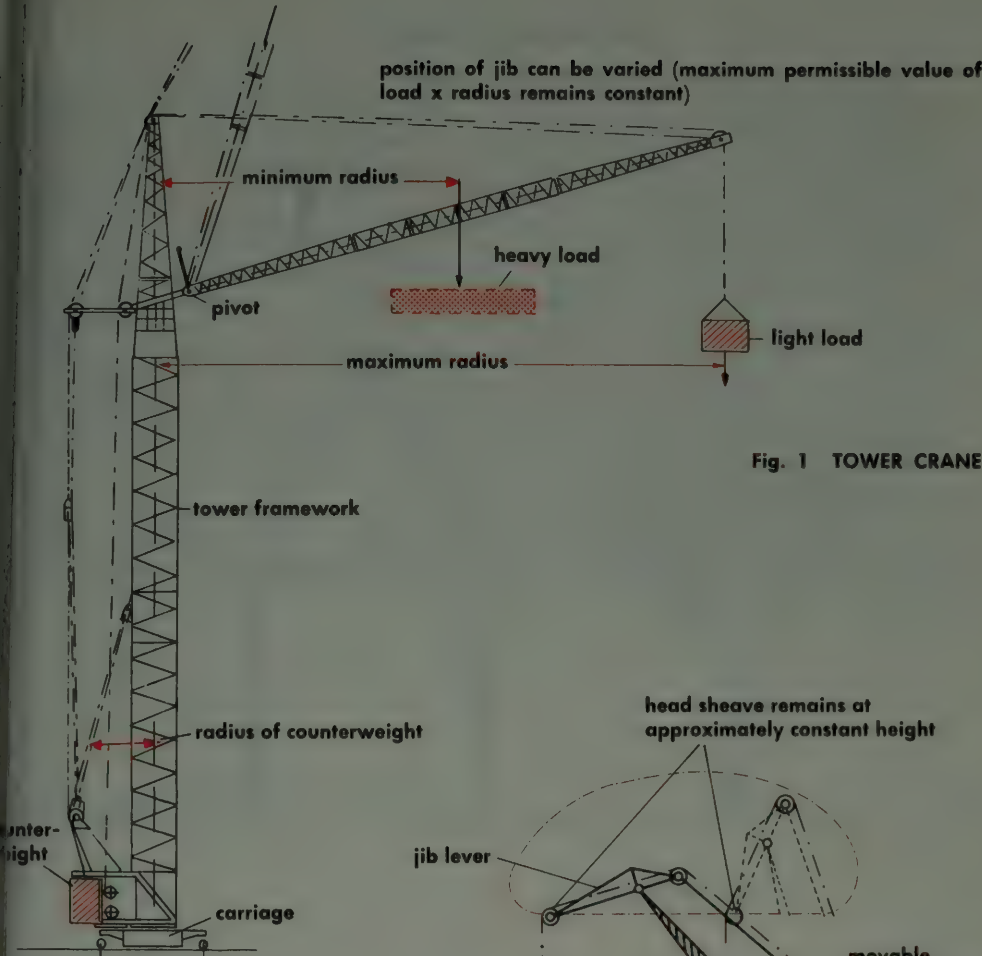


Fig. 1 TOWER CRANE

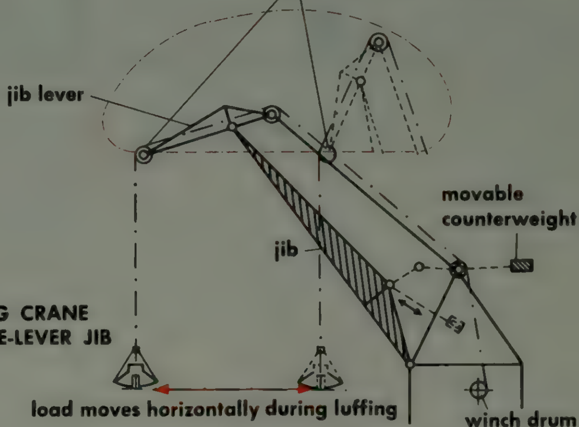


Fig. 2 LEVEL-LUFFING CRANE WITH DOUBLE-LEVER JIB

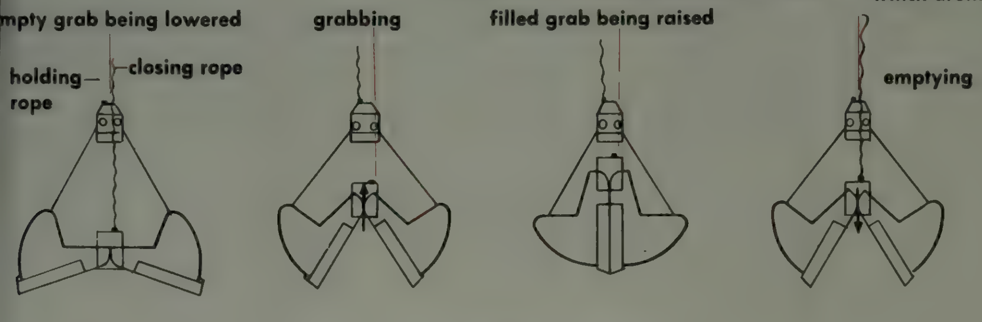


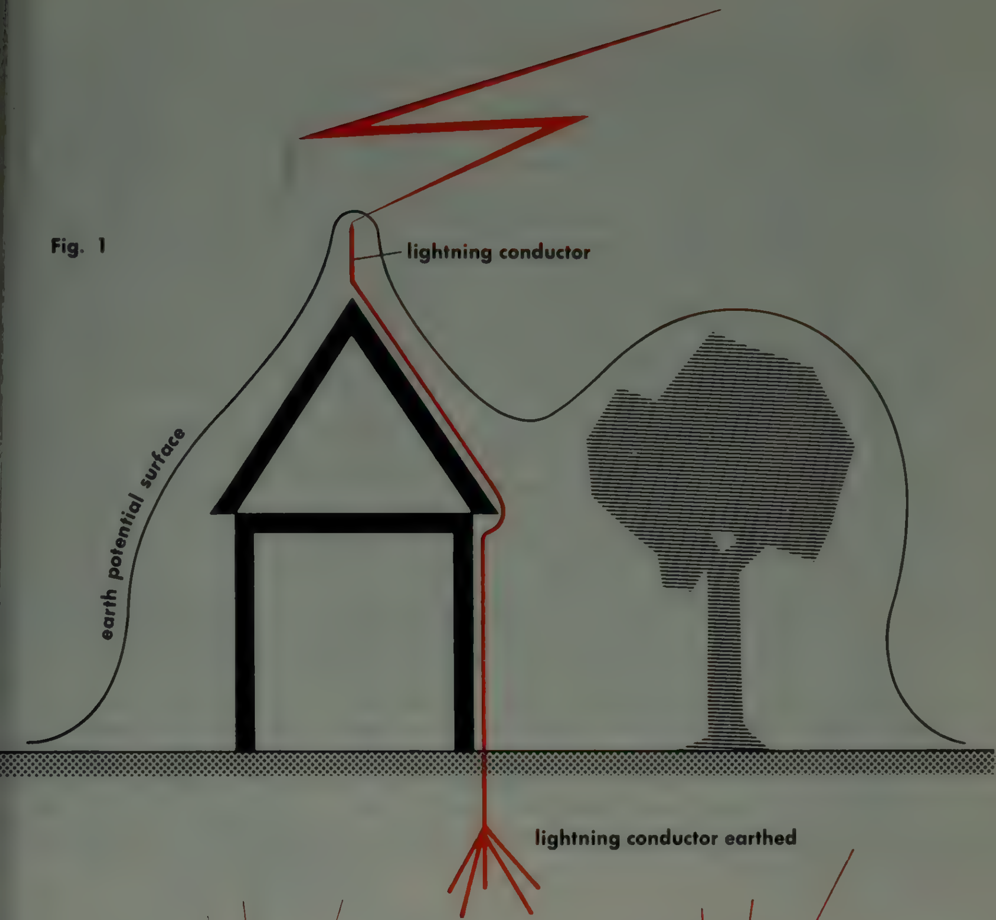
Fig. 3 HOW A GRAB IS OPERATED

LIGHTNING CONDUCTOR

A lightning conductor, or lightning rod, comprises a system of metal conducto whose purpose is to provide an easy path to earth for a lightning discharge strikir the highest points of a building. In this way the discharge takes place through th conductor instead of through the building, so that the latter is safeguarded froı damage. To perform this function, the lightning conductor must, of course, be pr perly earthed (grounded) (Fig. 1). The principle of the lightning conductor w discovered by Benjamin Franklin in the 18th century. The system consists of pointe air terminals (rods) mounted on the ridges of roofs, on chimneys, etc. Lightnin conductors do not prevent lightning strokes, but exert a local influence to dire strokes to the air terminals and thence safely to earth. To help them perform th function the air terminals are provided with sharp points. At these points the lin of force of the electric field are closely concentrated (Fig. 2), so that ionisation of th air around the points takes place, i.e., electrons (negatively charged particles) becor dislodged from atoms, with the result that the latter are left with a positive charg thereby making the air conductive to electricity and providing an easier path fc the lightning discharge (Fig. 3).

Steel-framed buildings generally need no special lightning protection, as the fram itself provides a suitable path for the lightning discharge to earth. Isolated building such as farm buildings, run a greater risk of being struck by lightning than building of similar size and height in built-up areas.

Fig. 1



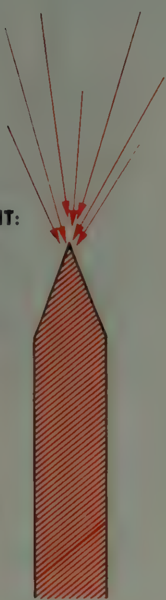


Fig. 2 **EFFECT OF SHARP POINT:** *electric lines of force are closely concentrated*

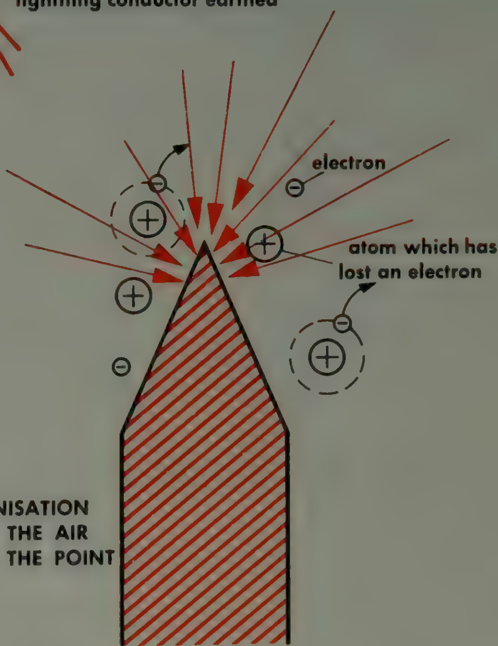
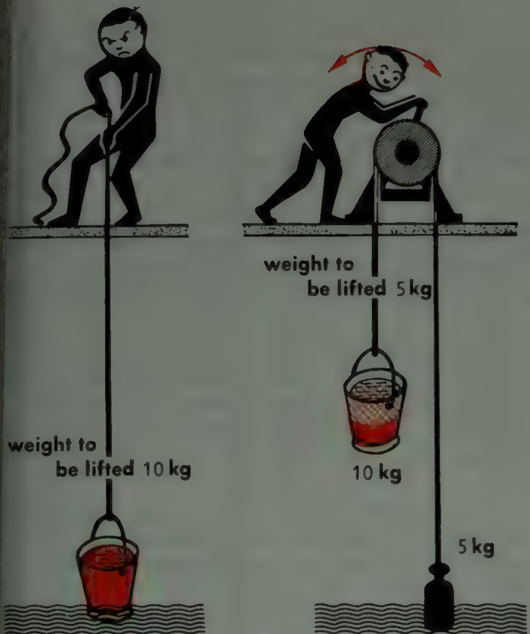


Fig. 3 **IONISATION OF THE AIR AT THE POINT**

A lift (or elevator, as it is called in the United States) is a power-operated device for lifting and lowering passengers from one level to another; it comprises a car which runs between guide rails and is suspended from steel hoisting ropes. The weight of the car and its load is approximately counterbalanced by a counterweight. The weight to be hoisted by the drive motor is therefore never the total weight of car and passengers but only the relatively small difference between the counterweight and the weight of the loaded car (this latter weight will, of course, vary to some extent, according to the number of passengers carried at any given time). The car is braked by means of electromagnets acting upon the drive shaft of the hoist pulley. The creep motion of the car before stopping at a floor is obtained by changing over to a lower speed of the motor. The switches which automatically effect this speed change are installed on the inside of the lift shaft. Similar automatic switches unlock the door or gate of the lift when it stops at a floor of the building. When the lift passes a floor without stopping, the device which actuates these switches is rendered temporarily inoperative.

Some lifts are equipped with a safety rope which runs in an endless loop round pulleys at the top and bottom of the shaft respectively (Fig. 2). This rope is secured to the lift car. In the event of fracture of the hoisting rope, the car will drop. This causes the pulleys of the safety rope to rotate more rapidly. A centrifugal governor (see page 214) connected to the top safety rope pulley then actuates a switch which sets the car safety device in operation. This causes powerful jaws to grip the guide rails and thus arrest the descent of the car. Various other safety devices are provided on modern lifts, including limit switches to prevent over-travel of the car, door interlocks to prevent the car from starting until the doors are securely closed, etc.

A type of lift still sometimes employed is the so-called paternoster lift (Fig. 3). This is a continuous elevating device for passengers, which consists essentially of two endless chains between which the cars are suspended and which are so arranged that the cars go up on one side and down on the other side. The two chains are driven by a motor installed at the top of the building. The cars travel continuously, but at low speed, so that there is sufficient time to step in and out of them at the successive floors. In modern buildings such devices have largely been superseded by escalators (see p. 270, vol. II)



g. 1 PRINCIPLE OF WEIGHT AND COUNTERWEIGHT

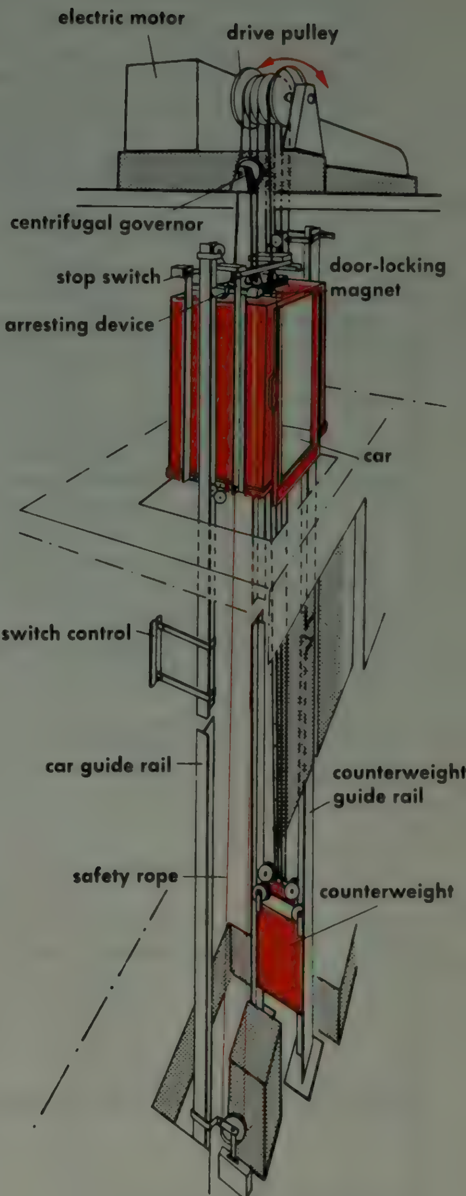


Fig. 2 LIFT INSTALLATION *(schematic)*

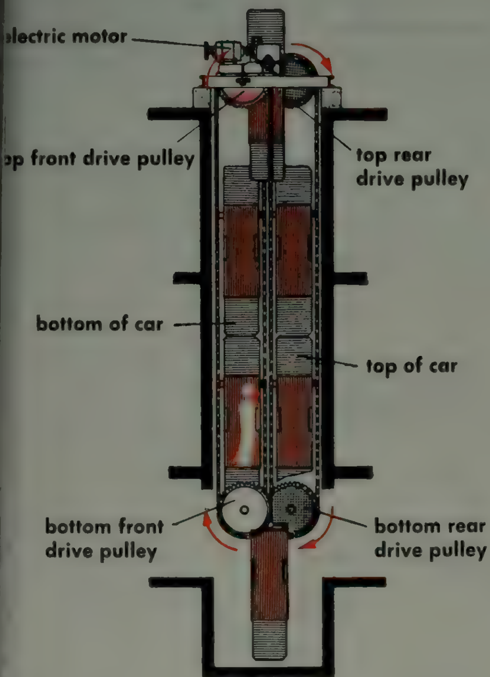


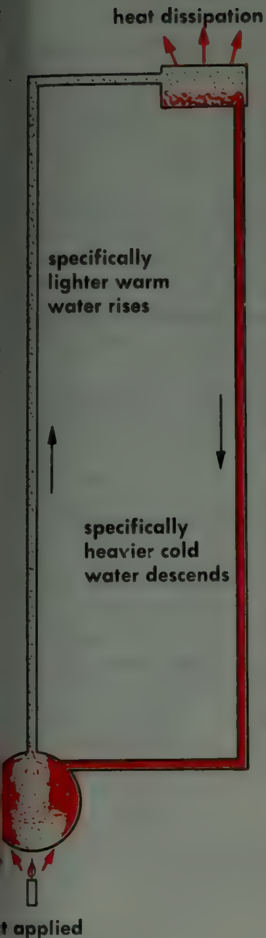
Fig. 3 PATERNOSTER LIFT *(schematic)*

In a heating system operated with hot water the water serves as the medium for carrying the heat to all parts of the building. Heating causes nearly all substances to expand. This is also true of water. A result of this phenomenon is that the specific gravity, or the weight per unit volume, of water decreases when its temperature is raised, i.e., warm water is specifically lighter than cold water. Fig. 1 illustrates the principle of hot-water heating. When the tank shown at the bottom left-hand corner of the system is heated, the hot water rises in the vertical pipe (because of its lower specific gravity, which makes it "float" in relation to the colder water), while cooler ("heavier") water descends in the right-hand vertical pipe and enters the tank where it is in turn heated and rises in the pipe on the left. In this way a constant circulation of water through the pipes and through the radiator in the top right-hand corner of the system is maintained.

In the domestic central heating system in Fig. 2 the hot water from the boiler rises through the flow pipe and circulates through the radiators. In these it gives off heat and therefore cools. It thus becomes specifically heavier and descends by gravity through the return pipes to the boiler, where it is heated and resumes its circulation through the system. For efficient circulation the boiler must obviously be located at the lowest point of the system.

In gravity systems the water is heated to about 90° C when it rises in the flow pipe. The radiators are so designed that the temperature of the water going down the return pipes is about 70° C. Venting the system, i.e., allowing entrapped air in the radiators to escape, is essential to ensure efficient functioning. A vent pipe may be installed at the top of the system; the radiators may be provided with individual venting cocks.

In the simplest form of gravity-circulation hot-water heating (Fig. 2) all the hot water ascends to roof level through the flow pipe and then flows down through return pipes and passes successively through the radiators on the various floors of the building. The disadvantage is that the radiators on the lower floors receive colder water than those on the upper floors and therefore have to be made larger in order to compensate for this. For this reason the so-called two-pipe system (Fig. 3) is preferred, in which each radiator has its own individual hot-water feed instead of (in the case of the radiators on the lower floor in Fig. 2) receiving water which has already passed through another radiator.



1 PRINCIPLE OF HOT-WATER HEATING *(single-pipe system)*

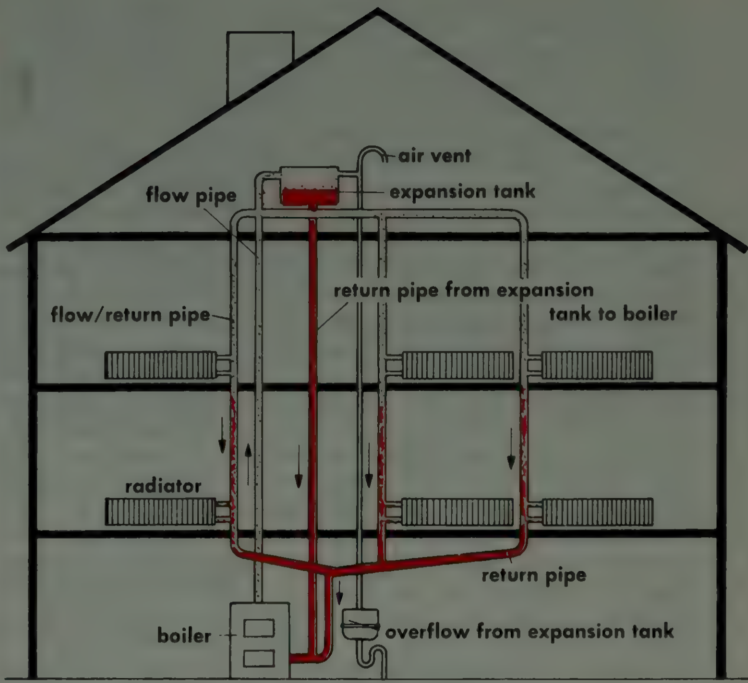


Fig. 2 OPEN GRAVITY HOT WATER HEATING *(single-pipe system)*

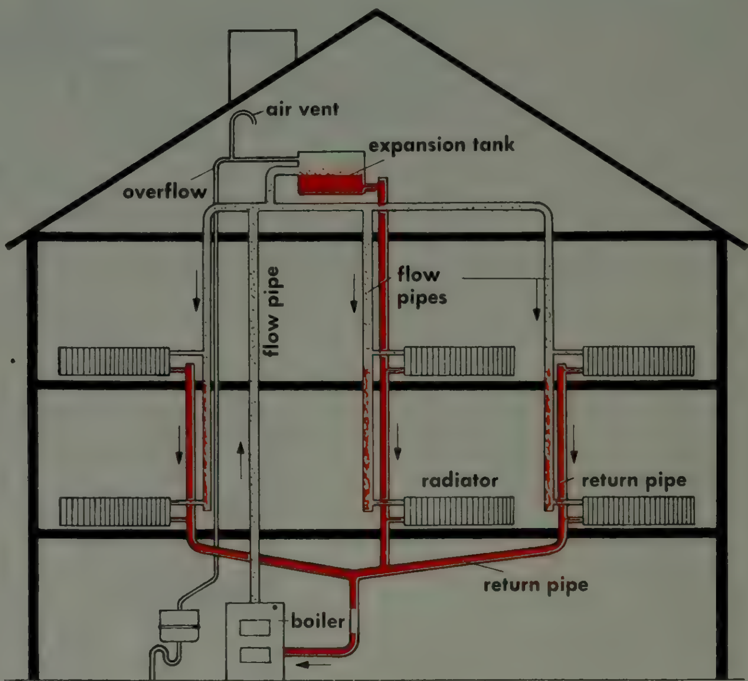


Fig. 3 OPEN GRAVITY HOT-WATER HEATING *(two-pipe system)*

Gravity systems rely entirely on the difference in specific gravity between hotter and colder water to produce the desired circulation and are especially suitable for small buildings and also for large buildings which are concentrated on plan. For large, sprawling complexes of buildings more efficient circulation can be achieved by means of an electrically driven centrifugal pump (cf. page 30). Pipes of smaller diameter than in gravity-circulation systems can be employed.

The pump may be installed either in the flow pipe (i.e., in the pipe extending from the boiler to the radiators) or in the return pipe. In the former arrangement the pump sucks hot water from the boiler and forces it through the pipes into the radiators, whence it returns to the boiler. When the pump is in the return pipe, it sucks the cooled water from the radiators and forces it into the boiler. In the simplest arrangement for a pumped-circulation hot-water heating system the radiators on successive floors are connected in series (single-pipe system, Fig. 1).

The two-pipe system with pumped circulation is similar in principle to the gravity system (see page 274). In addition, the pump may be installed in the flow pipe or in the return pipe (Fig. 3 and Fig. 2), as already stated. In the two-pipe system the hot water that flows into the radiators and the cooled water that flows out of them are conveyed through two different pipelines. The installations described here are known as low-pressure systems, i.e., they are in open communication with the external air through a vent pipe from the expansion tank.

Hot-water heating systems can be automatically controlled, i.e., the room temperature can be kept constant at any desired value by means of thermostats (see page 26). These devices are pre-set to the desired temperature, and when the actual temperature rises above, or falls below, this value, the thermostats transmit impulses which control the valves which in turn control the flow of hot water to the radiators or control the firing of the boiler itself by appropriately varying the fuel supply rate or (in the case of coal-fired systems) the combustion air rate. In the case of pumped circulation it is, alternatively, possible to control the hot-water flow by varying the speed of the pump.

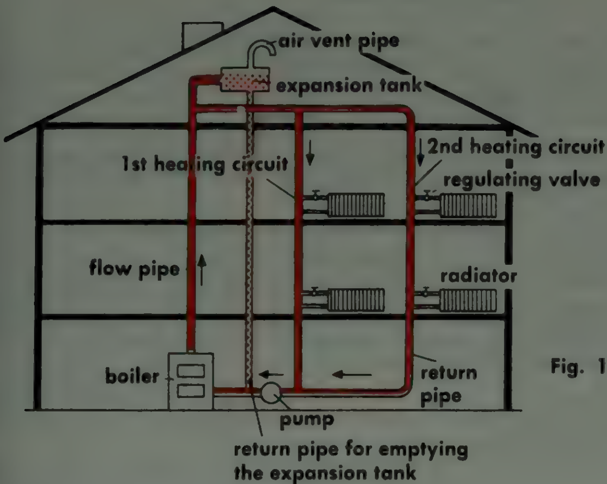


Fig. 1 SINGLE-PIPE PUMPED-CIRCULATION HEATING SYSTEM *with radiators connected in series (two circuits) and pump installed in the return pipe*

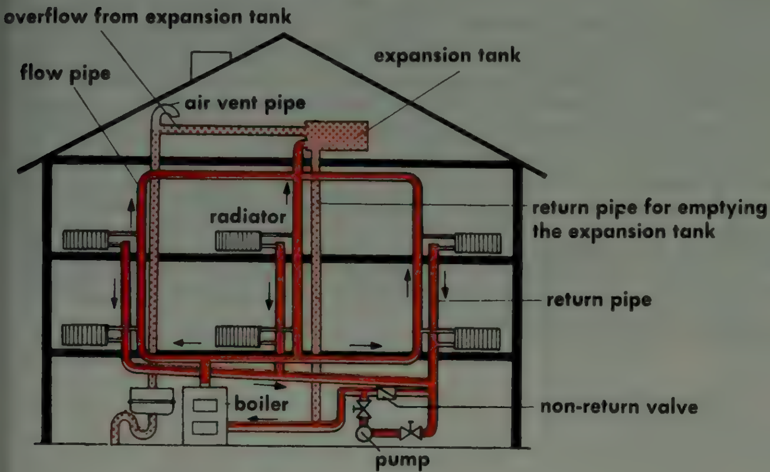


Fig. 2 TWO-PIPE SYSTEM *with pump installed in the return pipe*

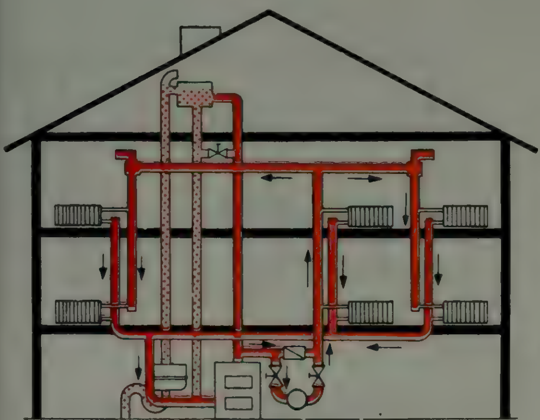


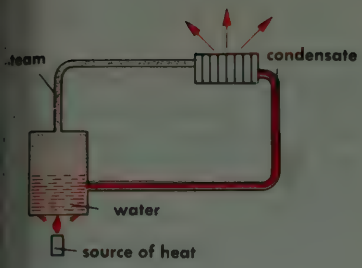
Fig. 3 TWO-PIPE SYSTEM *with pump installed in the flow pipe*

In this type of central heating the heat-conveying medium is not hot water, but steam. In the radiators the steam gives off its heat and condenses. It thus flows back as water to the boiler, where it is again turned into steam (Fig. 1). In houses, low-pressure steam heating systems are used; these usually have an excess pressure of about 0.5 atm. (7 lb./in^2.) in relation to the external air, with which the system is in communication through valves. The hot steam used has a temperature of around 105° C.

From the boiler the steam rises to the main distribution pipe, which is installed either in the basement (Fig. 2a) or in the top storey (Fig. 2b) of the building. From this pipe the steam flows into the radiators; the condensation water flows back to the boiler through the same pipe. This return flow pipe may be situated above the level of the water in the boiler. This arrangement may be described as "dry" condensate return. To prevent steam from also flowing back into the boiler, steam traps are installed in this pipe. In the second system the return flow pipe is situated below the boiler water level and is entirely filled with water and thus prevents the return flow of steam ("wet" condensate return).

In the two-pipe system, which is more often employed, the steam and the condensate flow in separate pipes. Here again bottom (Fig. 3a) or top distribution (Fig. 3b) may be employed. In both cases, however, the steam always enters the radiators from above and subsequently emerges (as condensate) from below. When the system is started up, and the pipes and radiators become filled with steam, the air that was present in them is forced to the bottom of the system (as steam is specifically lighter than air), where it must be discharged through vent cocks.

To prevent the pressure in the boiler from becoming excessively high, the steam pipe is connected by means of a branch pipe to a relief valve, which may be of the type illustrated in Fig. 4. Increasing steam pressure in the boiler forces the water down in pipe I and up in pipes II and III. If the pressure becomes excessive, the water in pipe I is forced down so far that steam can escape through pipe II and pipe IV. The water displaced from pipe II is collected in the tank and subsequently, when the pressure goes down, flows back into pipe III.



water turns to steam by heating

ig. 1 PRINCIPLE OF CIRCULATION IN A STEAM HEATING SYSTEM

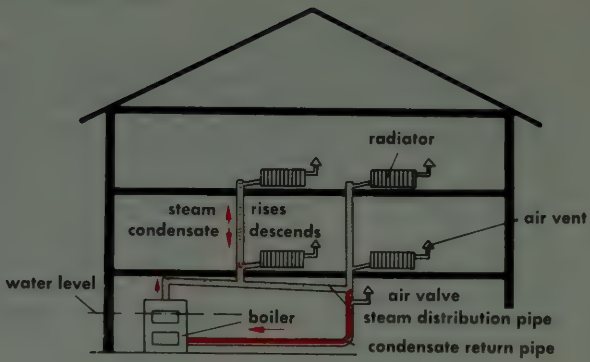


Fig. 2a SINGLE-PIPE LOW-PRESSURE SYSTEM WITH BOTTOM DISTRIBUTION AND "WET" CONDENSATE RETURN

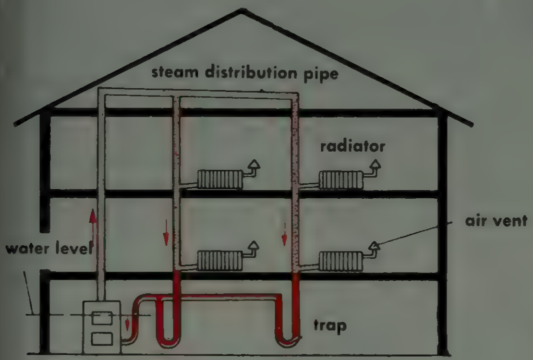


Fig. 2b SINGLE-PIPE LOW-PRESSURE SYSTEM WITH TOP DISTRIBUTION AND "DRY" CONDENSATE RETURN

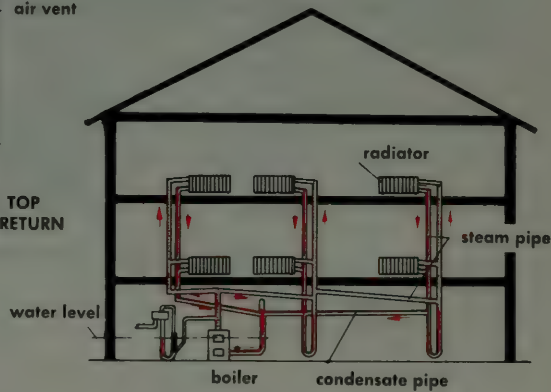
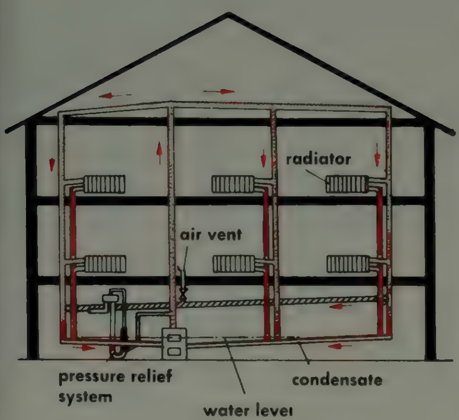


Fig. 3a TWO-PIPE LOW-PRESSURE SYSTEM WITH BOTTOM DISTRIBUTION AND "DRY" CONDENSATE RETURN



g. 3b TWO-PIPE LOW-PRESSURE SYSTEM WITH TOP DISTRIBUTION AND "WET" CONDENSATE RETURN

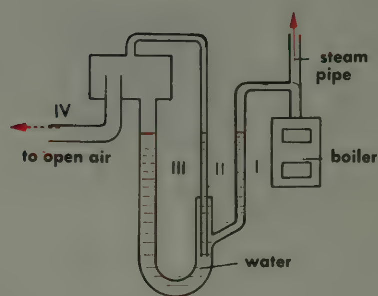


Fig. 4 PRESSURE RELIEF SYSTEM

Gas heating systems are usually fired with normal town gas. Like any other form of heating, gas heating functions on three physical principles: thermal radiation, convection, and conduction. Which of these three heat transmission effects predominates in any particular case will depend on the type of heating appliance.

One type of gas heater is the so-called infra-red radiator, whose heating action is based mainly on the emission of infra-red rays (these are heat rays and occur in the invisible range beyond the red end of the spectrum of visible light). These rays are produced by burning gas in special burners (Fig. 1a) whose heat is directed at panels of fireclay, steel fabric or some other suitable material, which are thus heated to glowing temperature and emit radiant heat. For heating halls and large rooms these radiant panels may be fitted to the ceiling (Fig. 1b). The radiation becomes more intense according as the panels are hotter.

So-called convection heaters represent a different heating principle. In this case the hot gases of combustion mainly give off its heat indirectly to the surrounding air. The most well known type of heating appliance functioning on this principle is the gas radiator, in which the burners are usually arranged in a row at the base. The hot combustion gases from the flames flow upwards through the radiator ribs which in turn give off the heat to the air. The latest form of construction of the gas-fired convection heater is the chimneyless type (Fig. 3), which can be mounted against any external wall. After giving off most of their heat, the combustion gases are discharged through a vent pipe straight into the open air. The significant feature of these convection heaters is that the hot gases are conducted through the appliance by the longest possible path, so as to enable them to part with as much of their heat as possible.

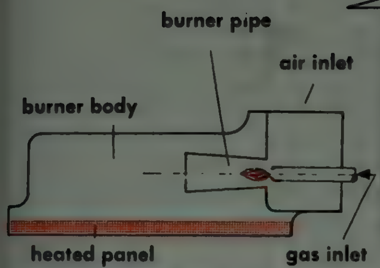


Fig. 1a INFRA-RED RADIATOR

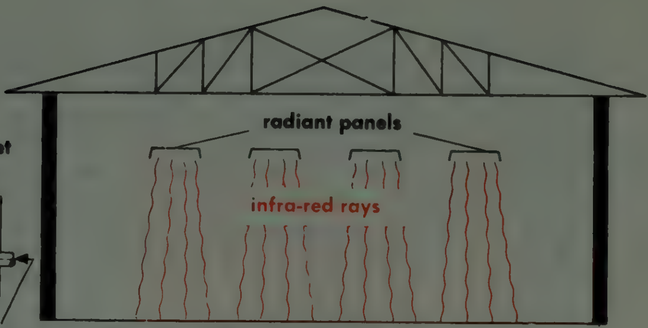


Fig. 1b BUILDING HEATED BY INFRA-RED RAYS

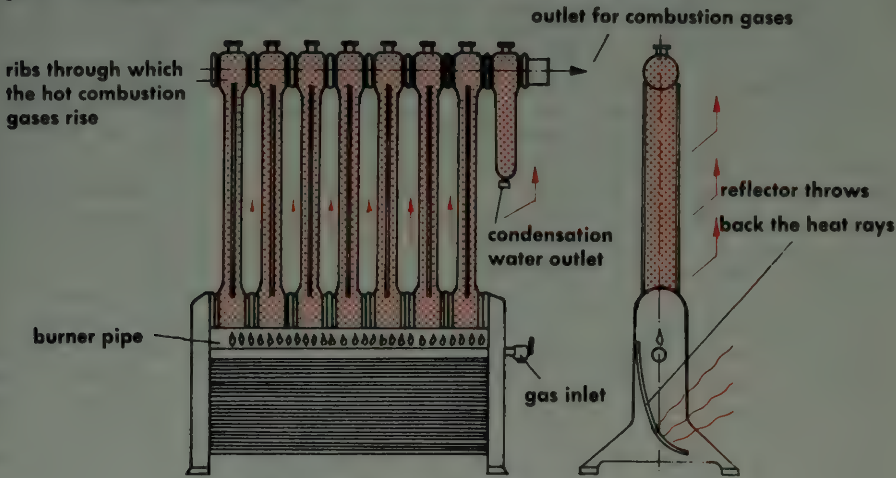


Fig. 2 GAS-FIRED CONVECTION HEATER

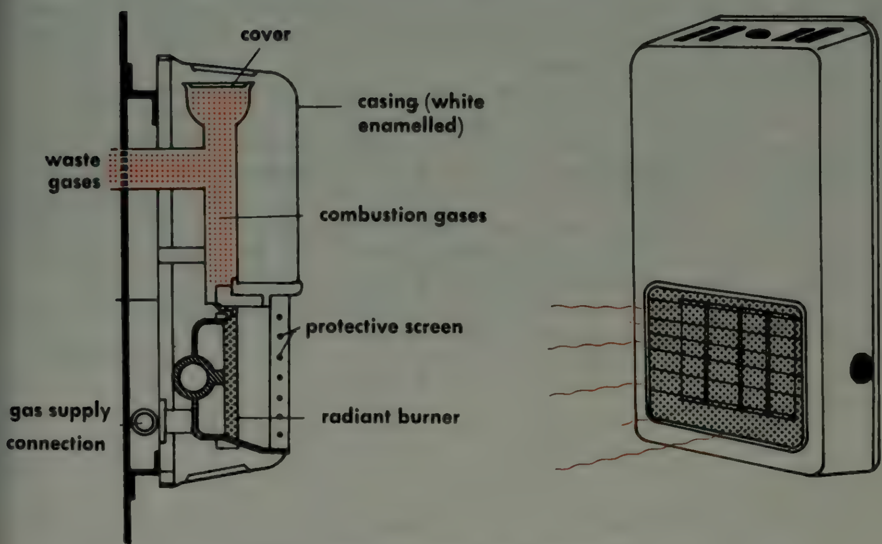


Fig. 3 CHIMNEYLESS GAS HEATER

The principle of the heat pump is similar to that of the compression refrigerator (see page 310). It has the same components: compressor, condenser, throttle valve and evaporator. But whereas the refrigerator extracts heat from a chamber by the evaporation of the refrigerant and thereby lowers the temperature, the heat pump supplies heat to a room by condensation of a heat-transfer medium. The phenomenon utilised for this purpose is that fluids which are under high pressure evaporate at a higher temperature than fluids under a lower pressure. In Fig. 1 two chambers are interconnected by a passage in which a vessel which completely seals it off can move from right to left and back. Atmospheric pressure exists in the left-hand chamber, while the right-hand chamber is under a much lower pressure (i.e., a partial vacuum). If energy is supplied to a quantity of water in the right-hand chamber, this water will evaporate ("boil") already at a temperature below 100° C. When a vessel containing water vapour formed by this process is shifted from the right-hand to to the left-hand chamber, the steam will at once condense at the higher pressure which prevails in the latter chamber. On condensing, it gives off heat. In this way a substantial proportion of the heat input in the right-hand chamber is recovered in the left-hand chamber (there can never be complete recovery, since some energy must be expended in pushing the vessel to the left, against the higher pressure).

In the heat pump (Fig. 2) the fluid heat-transfer medium—e.g., ammonia—is evaporated at low pressure in the evaporator. The heat needed for this may be obtained from various sources. For a small installation, the evaporator may, for example, merely be buried in the ground. The soil at all times contains sufficient heat to evaporate the heat-transfer medium. Other heat-transfer media that can be employed under suitable conditions are water and air. The work of transporting the medium from low to high pressure is done by the compressor. It draws in the vapour and compresses it to the desired higher pressure. Then, in a condenser, the steam is condensed at that higher pressure and gives off heat in doing so. This condensation is effected by passing the vapour through pipes with water flowing around them. The heat from the vapour of the condensing heat-transfer medium is thus transferred to this water, which is thereby heated to about 60°–70° C. and can be used as a second heat transfer medium to feed the radiators of a heating system. It cools in the radiators and flows back to the condenser, where it is heated up again. In so far as this circulating water is concerned, the condenser of the heat pump therefore performs the function of an ordinary central heating boiler. The vapour which condenses in the system is expanded to a lower pressure through a valve and is again evaporated in the evaporator.

By means of this cycle the heat-transfer medium within the heat pump is "pumped up" from a low temperature (e.g., the temperature of river water, about 10° C, say) to a higher temperature (high enough for a central heating system), the necessary work being done by the compressor. Heating systems based on this principle can be advantageous where electricity to drive the compressor is cheap. They are also sometimes installed in buildings where both cooling and heating are necessary, e.g., in dairies. The cellars can be cooled by the extraction of heat by the evaporator, while the heat evolved in the condenser heats the upper storeys.

Figs. 3a and 3b show a heat pump, which can be used for heating or cooling the room, depending on the setting of the four-way valve.

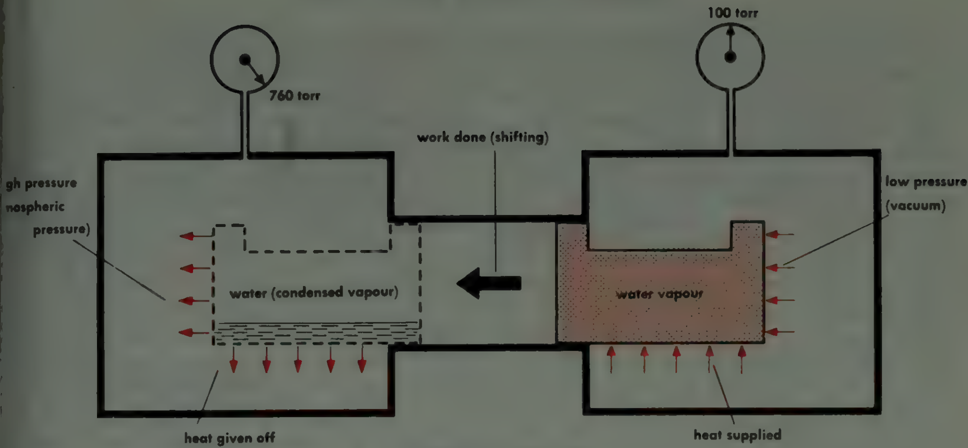


Fig. 1 PRINCIPLE OF HEAT PUMP

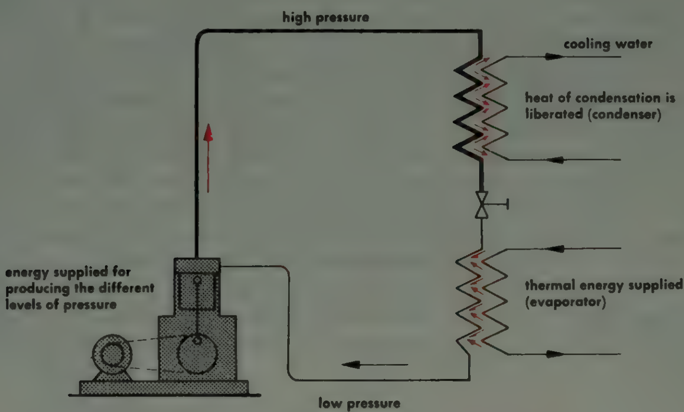


Fig. 2 DIAGRAM SHOWING OPERATION OF HEAT PUMP

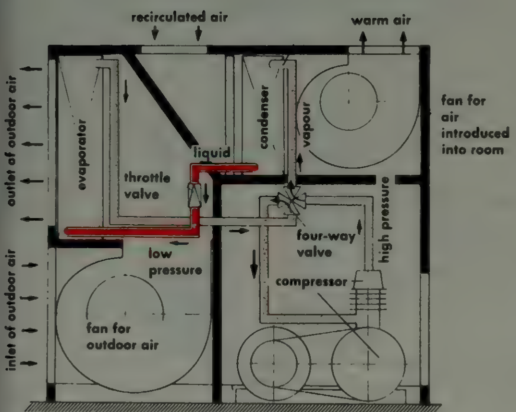


Fig. 3a HEATING

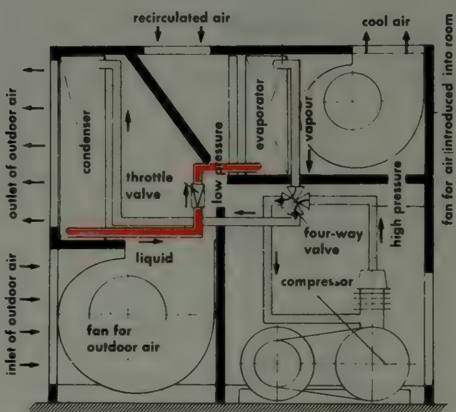


Fig. 3b COOLING

VALVES, COCKS AND TAPS

Devices known as valves, cocks, taps and faucets are used for controlling the flow of liquids and gases. "Valve" is the general term. The familiar water tap (or faucet) is technically therefore a valve. A plug cock (Fig. 1) is a simpler device in which the fluid passage is a hole in a rotatable plug which has a slightly tapered shape and is fitted in the body of the cock. Opening and closing the cock is done by rotating the plug through an angle of 90°. Some degree of flow control is possible by setting the plug of the cock in an intermediate position. Much more accurate control of the rate of flow of the fluid is provided by a valve, of which there are a number of different types. A feature common to most manually operated flow control valves is the stem (a screw spindle) surmounted by a handwheel. The fluid (gas or liquid) flows through an opening whose edge forms the valve seat. The ordinary domestic water tap illustrates the principle (Fig. 2). The stem is provided with a disc which is usually provided with a replaceable sealing washer to make the actual contact with the seat and thus stop the flow of water. To open the tap, the disc is raised by rotating the handwheel (or simple cross-bar handle) in the anti-clockwise direction, so that the stem is screwed out of the valve body. Clockwise rotation brings the valve disc into contact with the seat and thus closes the tap. In some valves the stem, instead of being provided with a disc, ends in a conical point which is inserted into the hole of the seat and closes it when the stem is screwed right home. This type is known as the needle valve. Another variant is the angle valve (Fig. 3). Larger valves used for controlling the flow of liquids in pipelines, water mains, etc. are often of the kind illustrated in Fig. 5. It is known as a gate valve or sluice valve. In this valve a wedge-shaped gate, actuated by a stem (with screw thread) and a handwheel, moves up and down. The gate bears against two seat faces to shut off the flow. The fluid flows through the gate valve in a straight line, so that the flow resistance is minimised. The valve stem must be suitably sealed at its point of entry into the valve body, in order to prevent leakage of the fluid. The seal is usually formed by a so-called stuffing box (Fig. 4), which is a cylindrical recess filled with packing which is compressed by a sleeve (known as a gland) to make a tight joint. The material inserted in the stuffing box is called packing. It may be a compressible material such as hemp or asbestos. The pressure exerted by the gland (screwed or bolted on) keeps the packing tightly pressed against the valve stem.

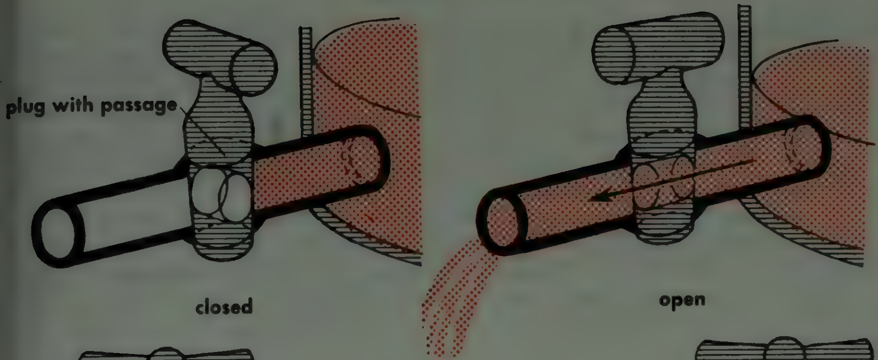


Fig. 1 PLUG COCK

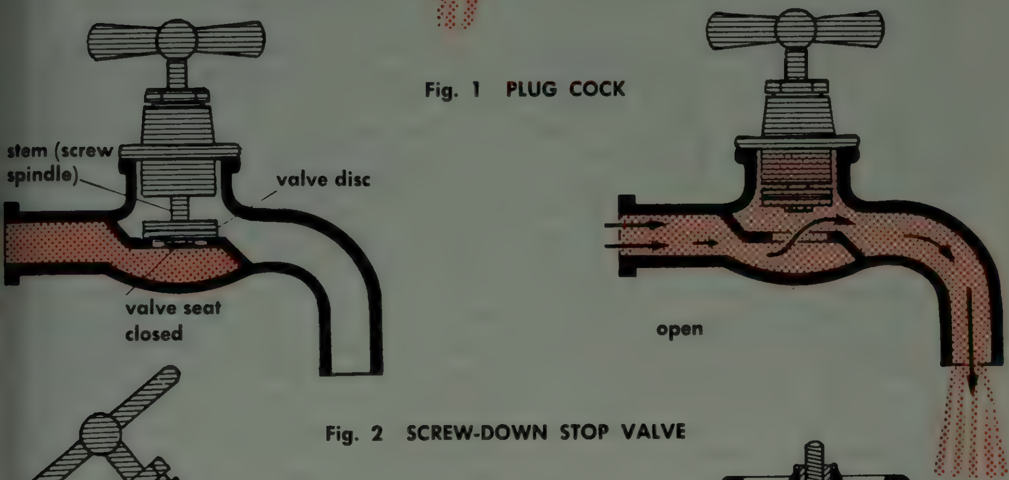


Fig. 2 SCREW-DOWN STOP VALVE

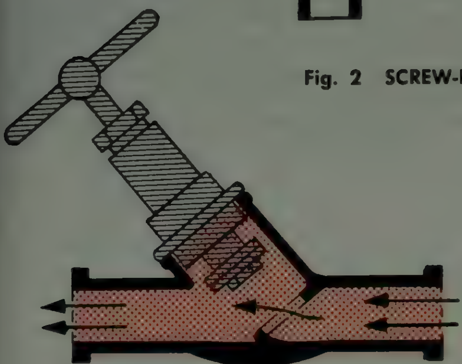


Fig. 3 ANGLE VALVE

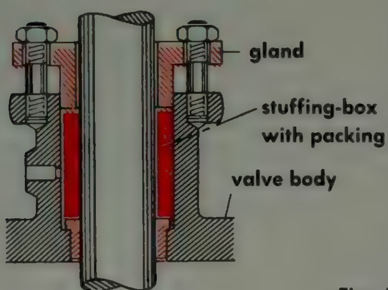


Fig. 4 STUFFING BOX

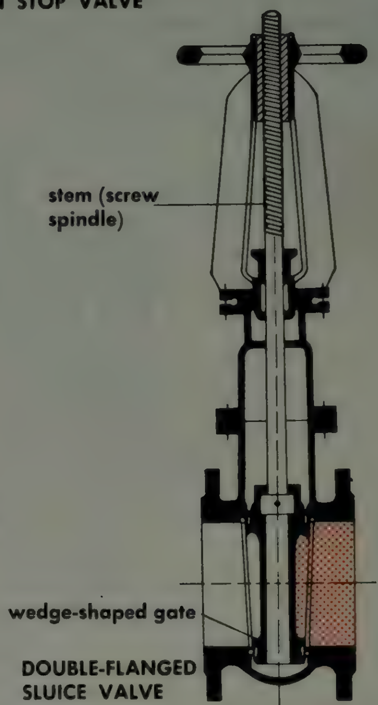


Fig. 5 DOUBLE-FLANGED SLUICE VALVE

The simplest kind of door lock is the so-called rim lock, whose metal case is screwed to the face of the door (Fig. 7). The lock is provided with a bolt which can be slid out or withdrawn by the action of the key. When the key is turned, it first presses the tumbler—which is kept engaged with the bolt by the pressure of a spring—upward and thus releases the bolt. The key bit (the lateral projection at or near the end of the key) then engages with the first notch on the underside of the bolt. Further rotation of the key causes the bolt to slide until the catch of the tumbler engages with the next notch on the top of the bolt. In simple locks, as on cupboard doors, this completes the locking action; but doors of rooms usually have double-action locks, in which the same operation is repeated and the bolt is moved along a further distance. Safeguard against unauthorised entry is provided merely by the individual shape of the bit (Fig. 2) and the corresponding hole into which it is inserted.

Greater safety is provided by a lock having not one, but several tumblers (Fig. 3). In such a lock the bolt is provided with a "stop" (a projecting pin). The tumblers are not notched at the top edge as in Fig. 1, but have their notches formed on the inside of a slot. The stop on the bolt engages with these notches. The undersides of the tumblers are variously shaped, and the key bit is provided with corresponding cuts and projections (Fig. 4). When the key is turned in the lock, the projections of the bit raise all the variously shaped tumblers an exact specified amount, whereby a clear passage for the stop through the slot is provided.

A further safeguard is obtained by the use of one or more wards. A ward may take the form of a ring (as in Fig. 5a) or a number of studs which are arranged in a circle around the centre of rotation of the key. The key bit is then additionally provided with grooves (Fig. 5b) which engage with the wards. If the key is not provided with such grooves in the correct position, it cannot be turned in the lock. The wards are fixed to the back plate of the lock or to a separate mounting plate (Fig. 6a); in that case the key has a divided bit (which fits over the plate), the grooves for the wards being arranged as shown in Fig. 6b. The lock illustrated in Fig. 7 has multiple lever tumblers similar to those in Fig. 3. For cylinder locks see page 288.

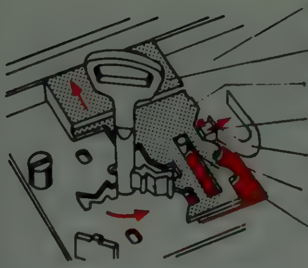


Fig. 1 LOCK WITH ONE TUMBLER


Fig. 2 KEY BITS

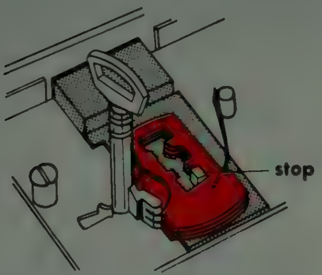


Fig. 3 LOCK WITH SEVERAL TUMBLERS


Fig. 4 KEYS FOR LOCKS WITH SEVERAL TUMBLERS

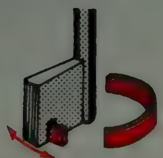


Fig. 5a WARD

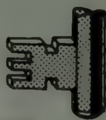


Fig. 5b KEY FOR LOCK PROVIDED WITH A WARD

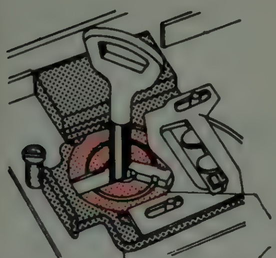


Fig. 6a WARDS FIXED TO SEPARATE MOUNTING PLATE

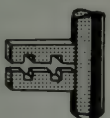


Fig. 6b KEY FOR LOCK IN FIG. 6A

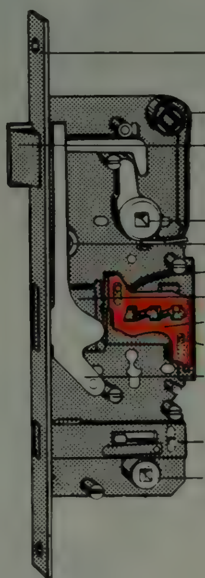


Fig. 7 RIM LOCK

The central feature of the cylinder lock is a rotatably mounted plug or cylinder. In the locked position (Fig. 1) a number of pin tumblers of different lengths and comprising an upper and lower segment are pressed down by springs to engage with holes in the cylinder, thereby preventing the latter from rotating. When the key is inserted into the lock (Fig. 2), the lower segments of the pin tumblers are raised by exactly the correct amount to bring their tops flush with the outer surface of the cylinder. As the two segments of each tumbler are separate, i.e., not interconnected, the cylinder is then free to rotate when the key is turned. The cylinder actuates the bolt, so that door can be opened. If the wrong key is inserted, it will not raise all or any of the lower tumbler segments to the correct height, and the cylinder cannot be rotated (Fig. 3).

Another and generally cheaper type of cylinder lock is the disc tumbler lock, in which the locking action is provided, not by segmented pins, but by movable discs which lock the cylinder.

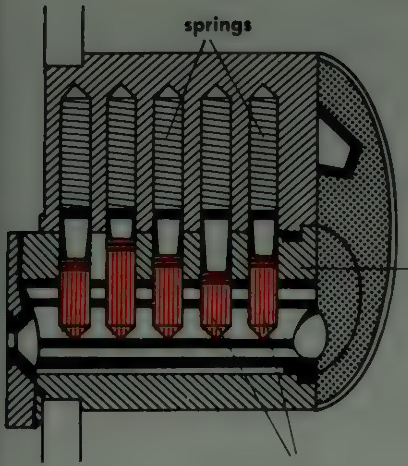


Fig. 1 LOCKED: PINS PREVENT PLUG FROM ROTATING

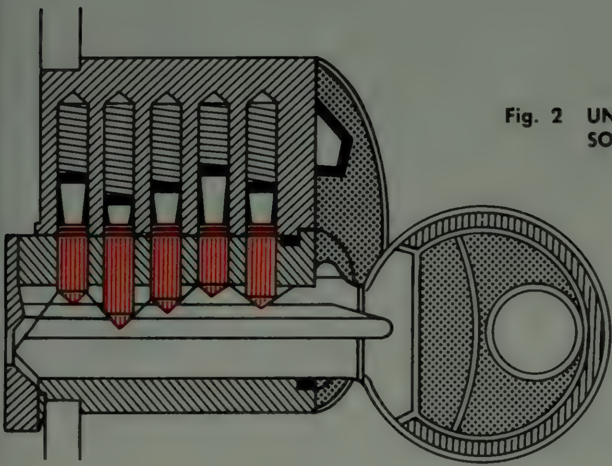


Fig. 2 UNLOCKED: KEY PUSHES UP THE PINS SO THAT PLUG CAN BE ROTATED

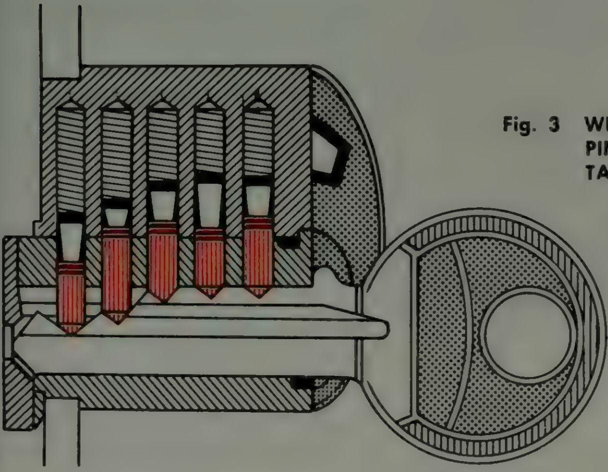


Fig. 3 WRONG KEY INSERTED: NOT ALL THE PINS ARE LIFTED A SUFFICIENT DISTANCE; PLUG CANNOT BE ROTATED

In an electric bell the to-and-fro movement of the hammer is produced by electromagnetic action. A two-pole electromagnet (comprising two cores interconnected by an iron yoke) is energised and attracts the armature to which the hammer is attached. At that instant the circuit is broken by the contact which is likewise attached to the armature; the electromagnet immediately releases the armature, which springs back, whereupon the contact re-establishes the circuit and thus causes the electromagnet to be energised again; and so on. This continues for as long as the push-button is pressed. A bell of this kind (or a buzzer, which is, in fact, nothing but an electric bell which has no hammer and no gong) can work on direct current or on low-frequency alternating current. For alternating current an alternative type of bell can be used, which requires no make-and-break contact (Fig. 2). The armature is polarised, i.e., it is permanent magnet. The electromagnet varies its polarity in the rhythm of the frequency of the current, so that the armature is periodically attracted and repelled. The hammer thus moves to and fro in time to the alternation of the current. Bells of these kinds are chiefly used in telephones.

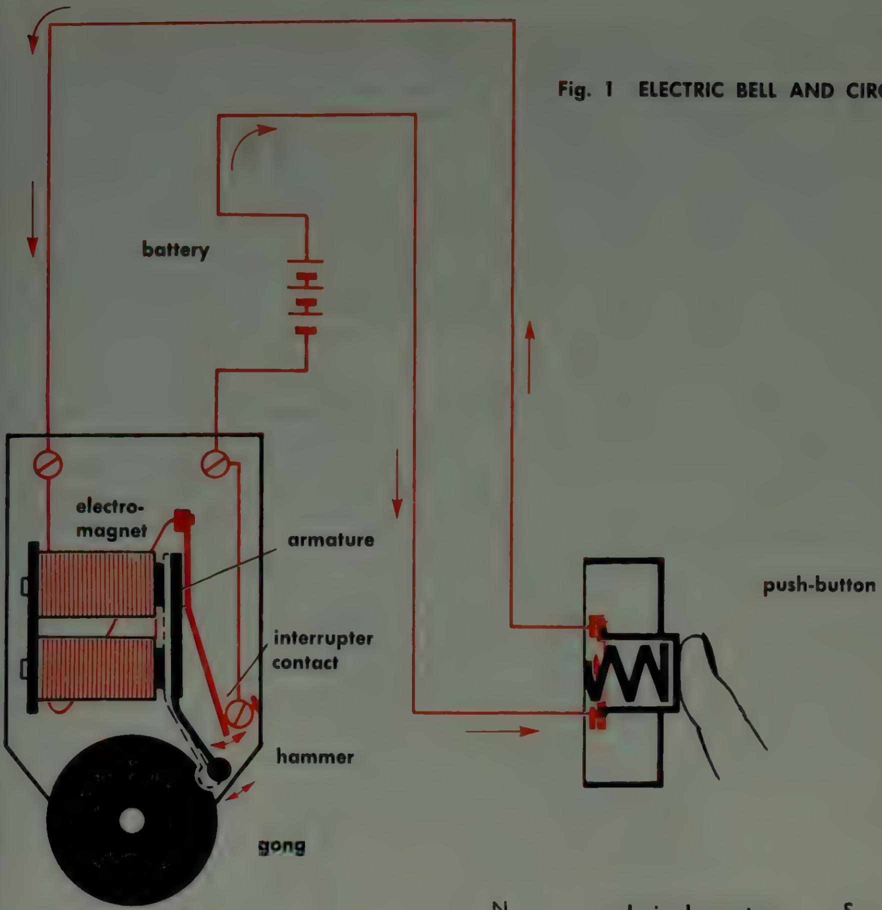


Fig. 1 ELECTRIC BELL AND CIRCUIT

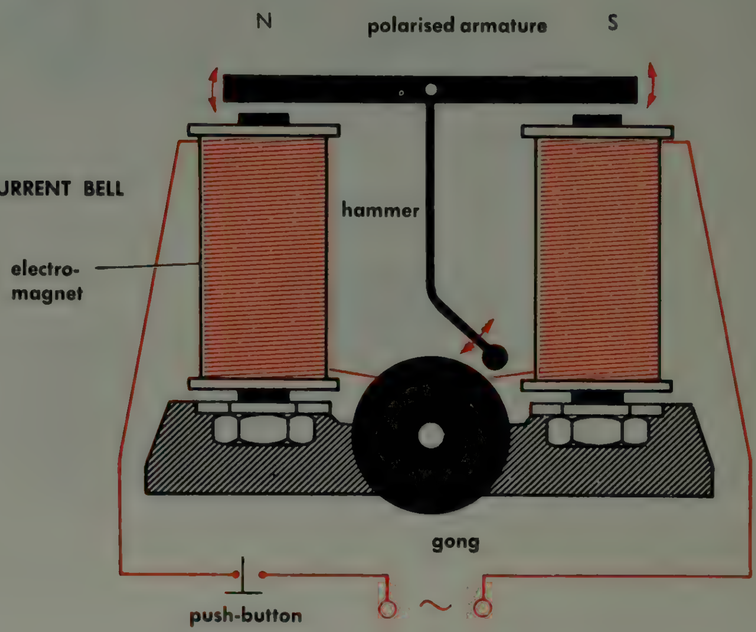


Fig. 2 ALTERNATING-CURRENT BELL

The door-closer is the familiar device that automatically closes an opened door. The simplest type (Fig. 1) works merely with a coil spring. It has the disadvantage that the door is shut rather violently. For this reason a more sophisticated device is used, which develops a slow closing movement. Door-closers of this kind operate on the principle of the shock absorber (p.232, vol. II): a piston moving in a cylinder forces oil through small openings. When the door is opened, the piston is withdrawn; and the oil in the cylinder opens a ball valve in a passage provided in the piston and flows into the space to the right of the piston (Fig. 2). When the door is released, it is closed by the action of the spring; the piston travels back to the right and forces the oil through the return passages (Fig. 3) into the space to the left of the piston. At first, the movement proceeds quite rapidly because the oil can flow back through two passages; but it becomes slower and slower, because the oil escape is progressively cut off by the advance of the piston.

The regulating screw provides the means of compensating for differences in the viscosity of the oil at different temperatures. In cold weather the oil is relatively thick. By unscrewing the regulating screw the flow passage through which the oil escapes is enlarged, so that the door can be closed just as briskly as in warmer weather, when the oil is thinner and therefore flows more easily. However, this adjustment is unnecessary if special silicone oils, which have an almost unvarying viscosity, are used in the door-closer.

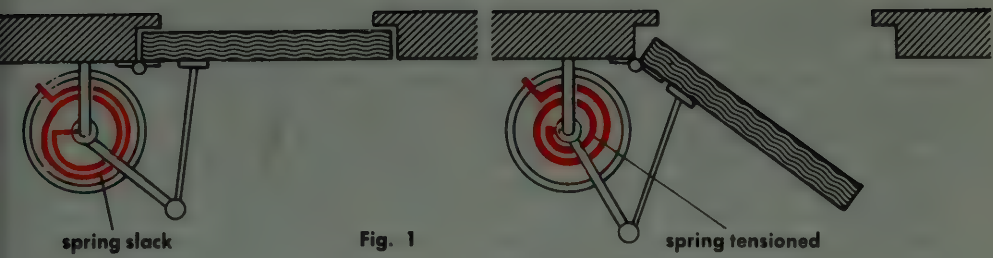


Fig. 1

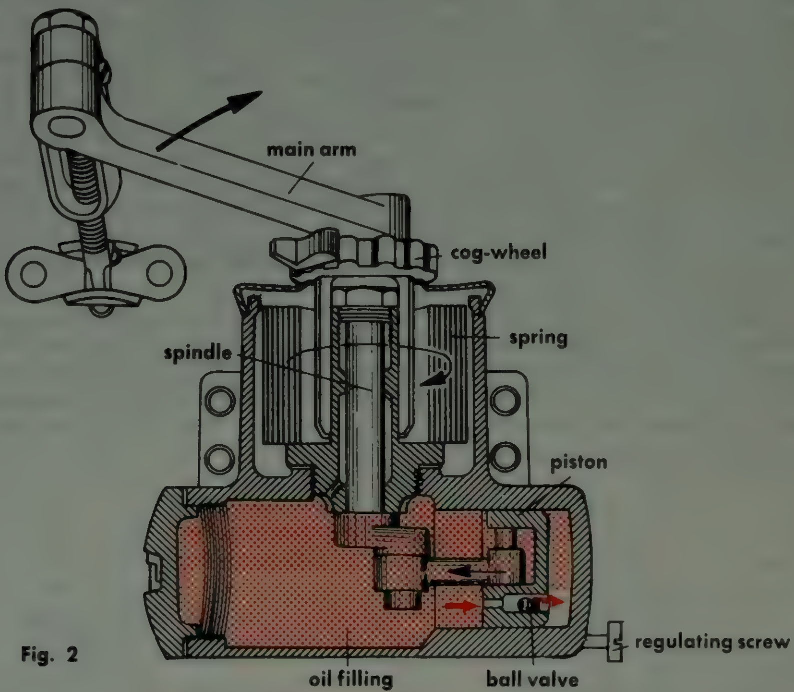


Fig. 2

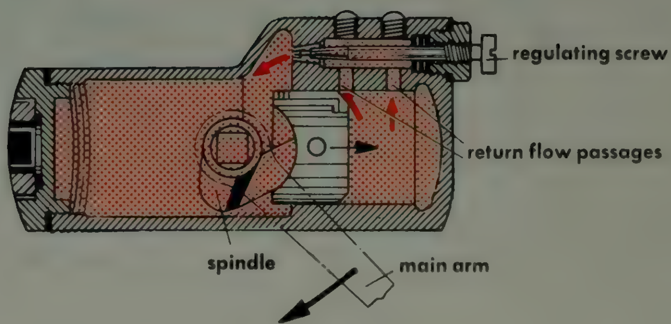


Fig. 3

When the chain attached to the lever of the flushing cistern is pulled (Fig. 2), the hollow iron bell-shaped unit rises and opens the passage to the flush pipe. As soon as water flows down this pipe, a vacuum is formed in the cavity of the bell and causes more water to flow from the cistern through the bell and down the pipe. The cavity inside the bell thus acts as a siphon (Fig. 1). When a vacuum is formed at *C* (by initially applied suction), water is drawn through the siphon tube. Once the flow has been started, it will continue. For the siphon to function, its outlet must always be below the level of the water in the tank. When the chain of the water closet has been briefly pulled and released, the bell falls back into position over the inlet of the flush pipe, but the flow of water down the pipe continues—thanks to the siphon effect—until the cistern has been drained. As the water level in the cistern goes down, the float descends and opens the water supply valve, so that the cistern fills up again. When the float has risen to a certain level, the inflowing water is cut off by the valve. The capacity of the flushing cistern is usually 2 gallons.

Fig. 3 illustrates another type of cistern. When the rod is briefly pulled up and then released, the water here, too, continues to flow until the cistern is drained. The rod is provided with a freely movable float which is prevented from floating to the surface of the water by two stops on the pull rod. When the rod is raised and the inlet of the flush pipe is opened, the closing pressure which is developed by the water column in the full tank is reduced. The buoyancy of the float predominates and keeps the pipe inlet open. Then the rod descends and the rubber valve disc is thrust against its seat by the inflowing water.

In some systems the flush pipe is connected to the water supply through a lever-operated (Fig. 4) or a push-button-operated valve (Fig. 5). In the former the flushing operation is initiated and terminated by hydraulic pressure equalisation which is effected by the composite valve system. In the push-button type, actuation of the button initiates the flow, which is subsequently likewise cut off by pressure equalisation and spring action.

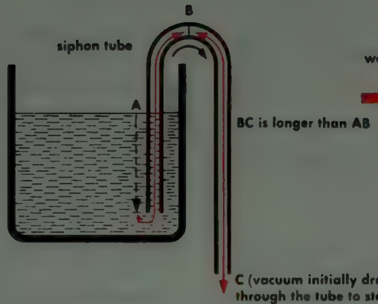


Fig. 1 PRINCIPLE OF SIPHON

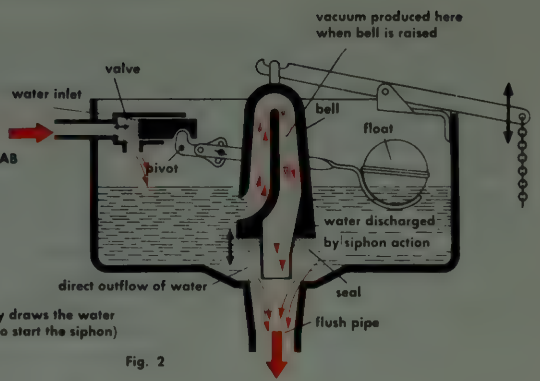


Fig. 2

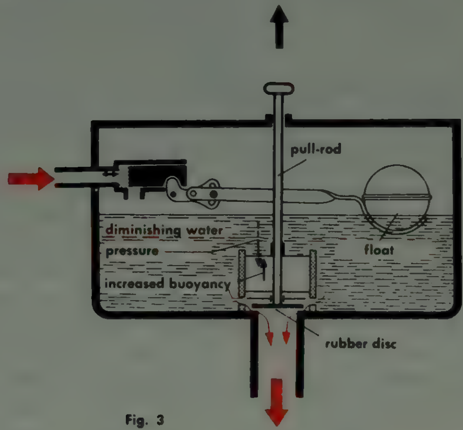


Fig. 3

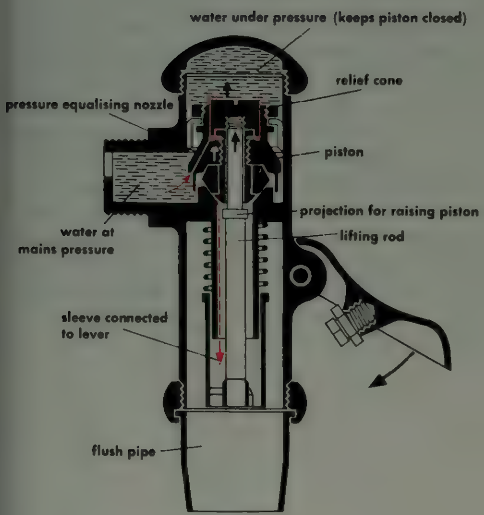


Fig. 4 LEVER-OPERATED FLUSHING SYSTEM

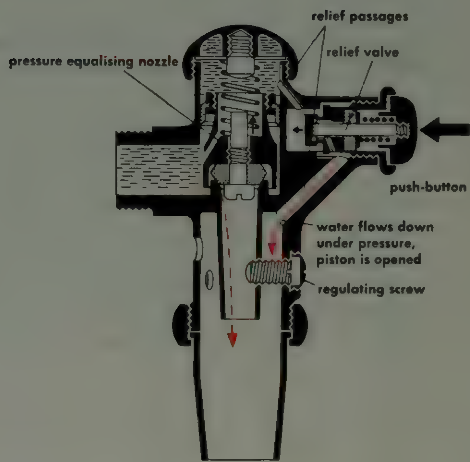


Fig. 5 PUSH-BUTTON-OPERATED FLUSHING SYSTEM

The function of an air-conditioning system is to keep the temperature and humidity of the air in rooms at values which provide a sense of comfort for human beings. The upper temperature limits of comfort are about 20° C and 25° C in winter and and summer respectively. At the same time, the relative humidity of the air must be between 35 and 70%. The relative humidity denotes the amount of water vapour actually present in the air as a percentage of the maximum amount that could be present at that particular temperature, i.e., if the air were saturated with moisture. It is therefore the ratio of the pressure of water vapour present to the saturated vapour pressure at the same temperature.

The central feature of an air-conditioning system is the air-conditioning plant (Fig. 1). Fresh air, together with a proportion of the air returned from the air-conditioned rooms (recirculated air) is drawn into a mixing chamber, the relative quantities of fresh air and recirculated air being controlled by valves. This air is cleaned by means of filters. These usually have glass-wool filter elements in which dust is retained.

After filtering, the air is preheated. This is done by means of heating pipes through which steam or hot water is passed and which are provided with fins which serve to increase their heat-exchange surface area. The air to be preheated flows along these fins and absorbs heat from them. Excess moisture is removed from a portion of the air by cooling. The warmer the air is, the more water vapour it can absorb. Conversely, when air with a certain moisture content is cooled, water condenses in the form of myriads of tiny droplets, which appear as fog. In the cooler of the air-conditioning plant the moisture is thus precipitated. The moisture content and temperature of the air emerging from the cooler are determined by the temperature of the cooler. This air is then mixed with the air coming straight from the preheater, so as to obtain an air mixture of the desired temperature. If the moisture content of the mixture is too low, finely atomised water is added by spray nozzles. This causes some cooling of the air, and for this reason the air is passed through a reheater which is essentially similar to the preheater and which gives the air its desired final temperature. Behind the reheater is the fan which forces the conditioned air through ducts to the various parts of the building. The constant supply of air to the rooms produces a slight excess pressure in them, which causes the exhaust air to flow back through return ducts. Some of this exhaust air is discharged to the open air, while a certain portion (the recirculated air) is returned to the conditioning plant. In smaller air-conditioning units the air sometimes is divided into two streams, one of which is heated and the other cooled (Fig. 2). In each room these two air streams are mixed in the proportions to produce the temperature desired by the occupants of that particular room.

The system described in the foregoing is what is known as a centralised air-conditioning system, i.e., all the air is treated in a central conditioning plant and conveyed through ducts to the rooms. There are many variants of the centralised system. Decentralised systems operate either in conjunction with a central conditioning plant which treats only the proportion of outside air that is introduced into the system, or they may be fully decentralised, using only self-contained cabinet or box-type air-conditioning units in the various rooms.

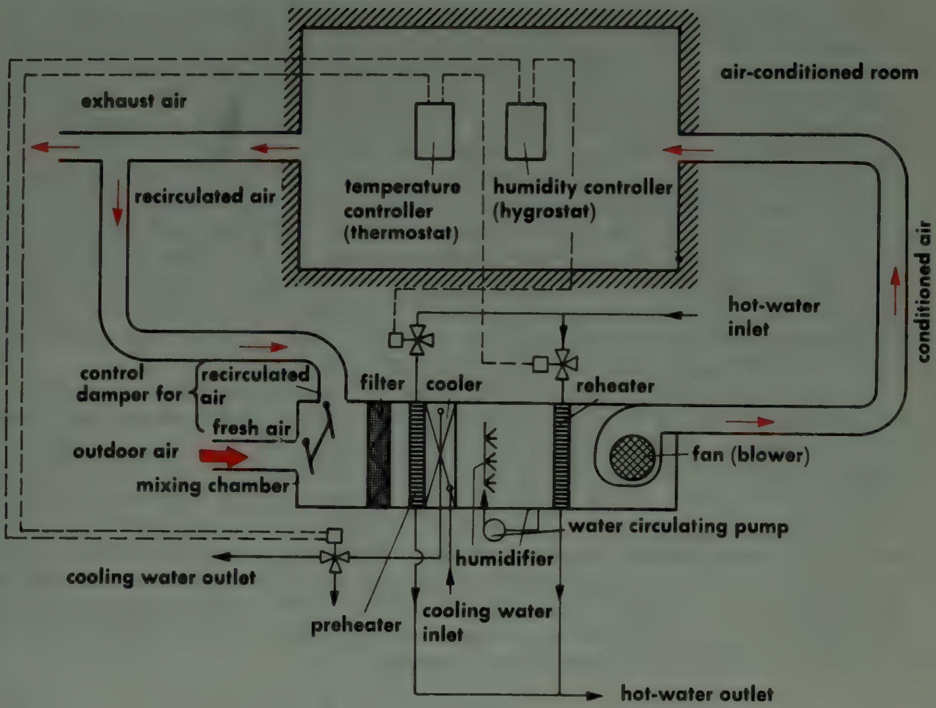


Fig. 1 CENTRALISED AIR CONDITIONING SYSTEM

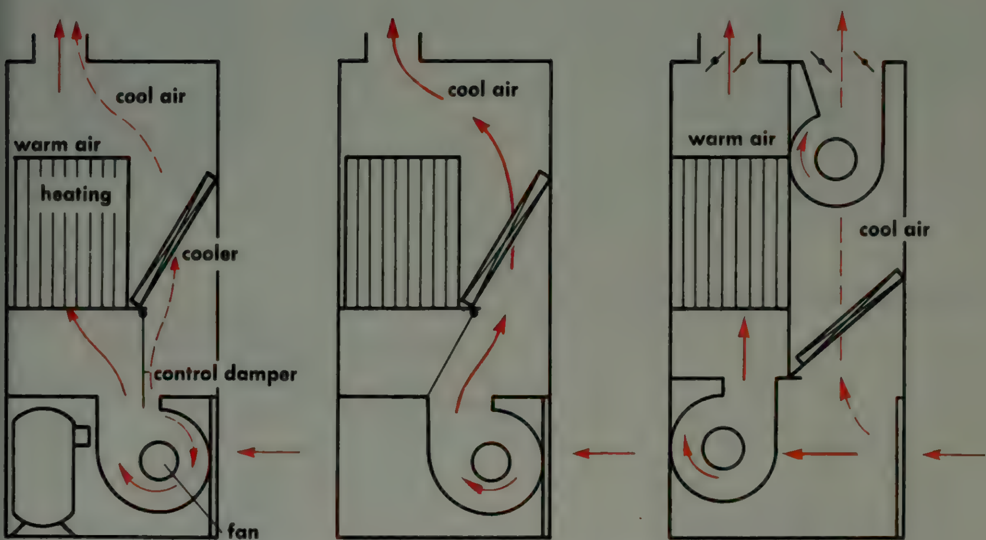


Fig. 2 SELF-CONTAINED AIR-CONDITIONING UNITS

Mechanical devices for moving air or other gases and operating on the centrifugal principle—as also applied in centrifugal pumps (cf. page 30)—are sometimes classified as fans (high rates of delivery at low pressure), blowers (high rates of delivery at medium pressure), and compressors (high-pressure delivery). In semi-technical language, however, the term "fan" is more often applied to any propeller-type device which imparts motion and acceleration to air (Fig. 2).

The operating principle is the same for all three classes of centrifugal machine. The air is drawn in at the centre of the casing (Fig. 1a) by the rotating impeller which is driven by an electric motor and contains a number of passages arranged in a spiral pattern (Fig. 3). On flowing through these passages the air is given an acceleration and emerges under pressure from the spiral casing (volute) of the fan (Fig. 1b). To obtain higher delivery pressures, a number of such impellers, mounted on the same shaft, can be installed one behind the other (in series), whereby the desired high pressure is achieved in several successive stages (Fig. 4).

Air can also be compressed by rotary methods based on positive displacement. Fig. 5 illustrated the functioning of a rotary compressor of this kind. A cylindrical rotor is disposed eccentrically in a cylindrical casing. The rotor is provided with approximately radial plates which are movably inserted in slots. As a result of the high speed of rotation, these plates are flung outwards by the centrifugal force and are thus passed against the inside of the casing. Because of the eccentricity of the rotor, the compartments between these plates become alternately larger and smaller, so that the air drawn into the compartments from the inlet pipe is compressed and discharged under pressure from the outlet pipe. In the Roots blower (Fig. 6) two mating lobed impellers, driven by two gear wheels which are in mesh with each other, revolve in opposite directions within a casing.

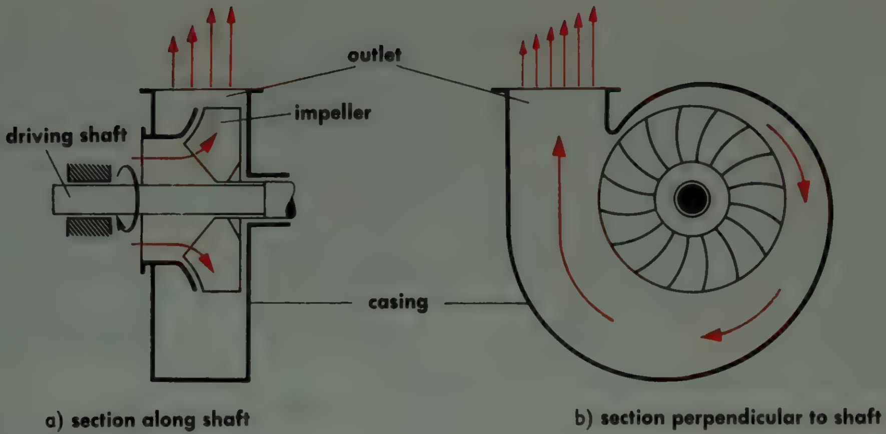


Fig. 1 CENTRIFUGAL COMPRESSOR

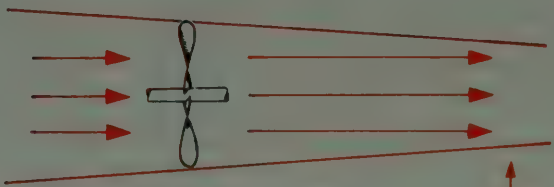


Fig. 2 FAN *(schematic)*

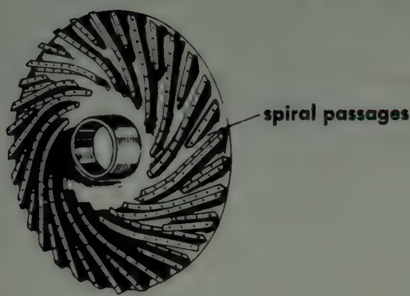


Fig. 3 IMPELLER

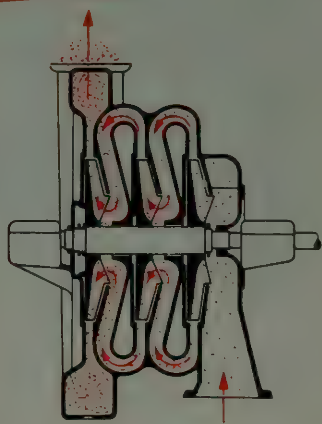


Fig. 4 MULTI-STAGE COMPRESSOR

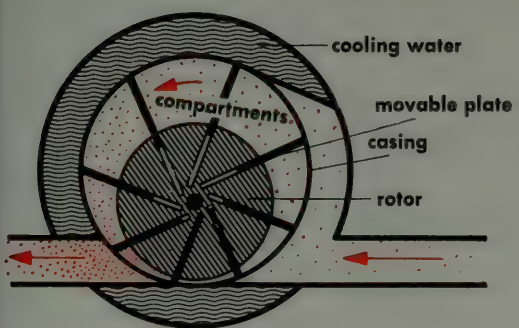


Fig. 5 ROTARY COMPRESSOR

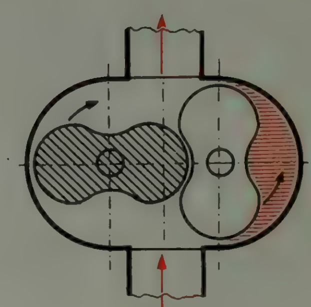


Fig. 6 ROOTS BLOWER

The most familiar type is the coal-burning slow-combustion stove (Fig. 1). It has a combustion chamber, lined with fireclay, which is charged with coal. The air for combustion enters the combustion chamber from below (by the action of natural draught). In another type of stove the air is drawn through the fuel in the downward direction (Fig. 2). Combustion takes place more particularly at the bottom of the combustion chamber; the narrower portion above it serves as a feed hopper from which fresh coal is supplied as the coal on the grate burns away (base-burner stove).

For better heat utilisation the combustion gases must not be directly discharged up the chimney, but should be compelled to make a detour so as to give off more of its heat. For this purpose the gases are made to pass through a system of ducts called flues. An efficiently designed stove embodying this principle can achieve as much as 80 per cent heat utilisation. Even so, the temperature of the gases discharged up the chimney is still about 200° C; the surface of the stove has approximately the same temperature when it is operating at full capacity.

A further development is represented by the so-called tile stove (Figs. 3a and 3b), which consist of an iron stove surrounded by a kind of jacket or casing constructed of tiles or some other ceramic material. The flue gases additionally give off heat to this casing, which stores up the heat and gradually gives it off to the air in the room.

Oil-fired stoves (Figs. 4 and 5) operate as follows: The oil from the supply tank flows into a small control tank which contains a float-operated valve that cuts off the inflow when the oil in the control tank has reached a certain level. In this way an approximately constant oil feed pressure to the burner is ensured. From the control tank the oil flows through a control valve which can be set to deliver the oil to the burner at a certain rate. By varying the valve setting, the heat output of the stove can be regulated. The burner is usually of the pot type, in which the oil first vaporises as a result of coming into contact with the hot walls of the combustion chamber. Combustion air flows into the chamber through holes, and the oil vapour burns continuously. The hot gases of combustion flow upwards and heat the combustion chamber walls. These in turn give off their heat to the air in the room. Some stoves are equipped with small fans to assist the flow of combustion air, thus making this largely independent of chimney draught. The fuel used is a relatively light fuel oil. The hot gases discharged up the chimney have a temperature of 300°–400° C.

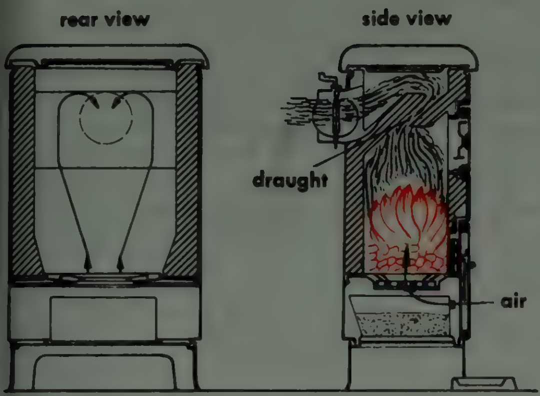


Fig. 1 COAL-BURNING SLOW-COMBUSTION STOVE

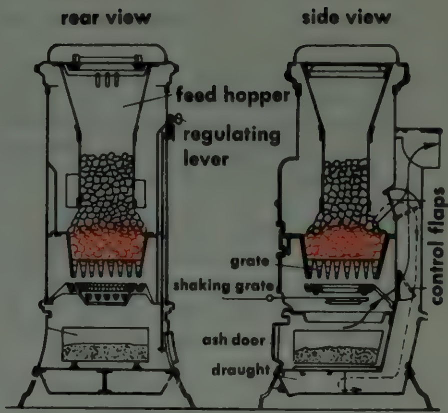


Fig. 2 BASE-BURNER STOVE

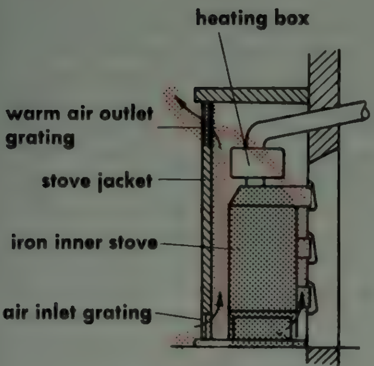


Fig. 3a TILE STOVE

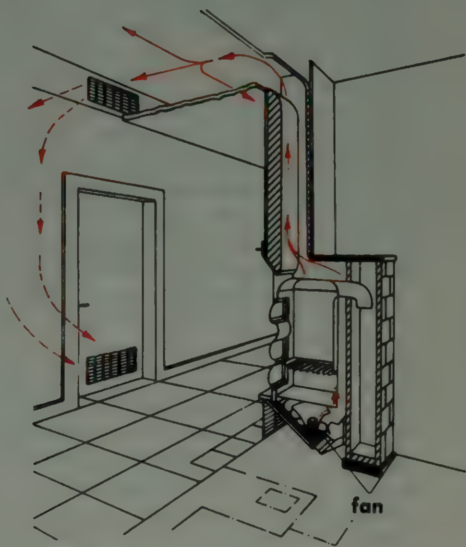


Fig. 3b TILE STOVE WITH FAN-ASSISTED AIR CIRCULATION

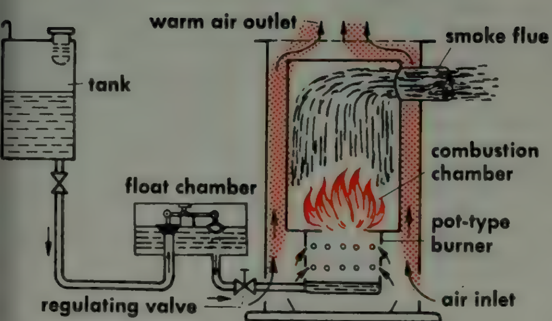


Fig. 4 OIL STOVE WITH VAPORISING BURNER

a) fan mounted under the burner
b) fan mounted on one side of the burner

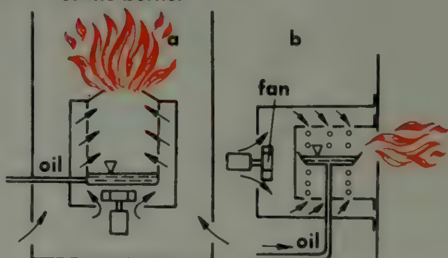


Fig. 5 VAPORISING BURNER WITH FAN

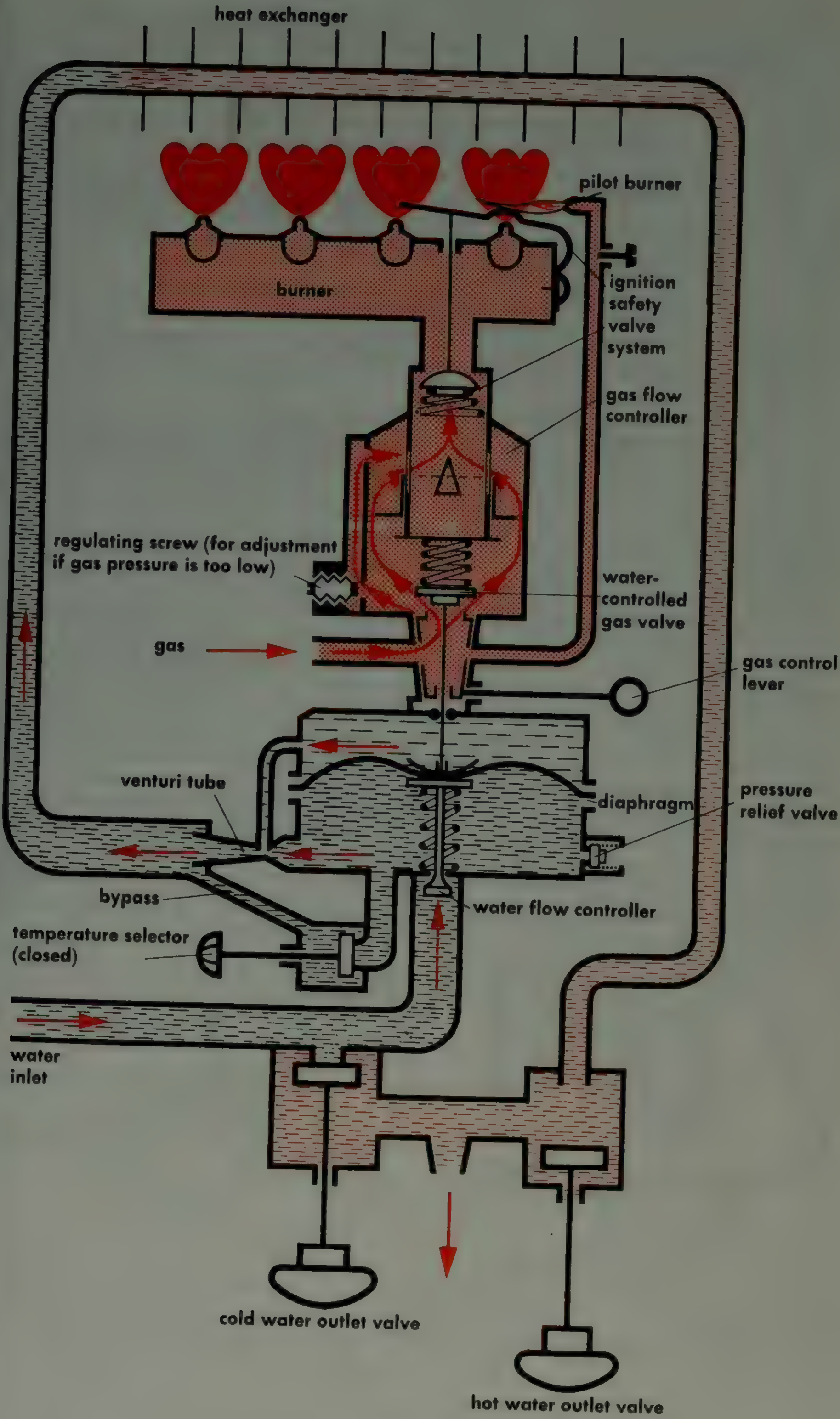
GAS-FIRED WATER HEATER (GEYSER)

In the so-called "instantaneous" gas-fired water heater (flow heater or geyser) the water flows continuously through the appliance, i.e., there is no storage vessel. The mode of functioning of such a heater is illustrated by the accompanying diagram.

The central feature is the heat exchanger through which the water flows and which is surrounded by the hot combustion gases produced by the burners. In the heat exchanger, which has a large heat transfer surface area, a large proportion of the heat contained in these gases is transferred to the water. The heat transfer is chiefly governed by the difference in temperature between the water to be heated and the hot gases and by the flow velocity of the water through the heat exchanger.

The pilot flame is kept alight and instantly lights the burner when the main gas supply is turned on. The gas flow controller keeps the rate of supply of gas to the burner constant and protects the applicance from overloading. The gas valve in the gas flow controller is itself controlled by the water flow acting through the agency of a diaphragm and a venturi tube (see page 260). The water flow controller keeps the water flowing at the desired pre-set rate by throttling down the mains pressure. With this device the flow of water through the heat exchanger can be regulated, and the final temperature of the hot water can thus be varied.

To get hot water, it is necessary first to move the gas control lever to the igniting position and light the pilot flame (assuming that this has not been done already). The heat of this flame causes the end of a bimetallic spring (see page 26) to move downwards, causing the ignition safety valve of the burner to open. When the gas control lever is moved farther, to the "on" position, gas is admitted under the water-controlled gas valve (which is still closed, however). Now when the hot-water outlet valve is opened, the flowing water causes an excess pressure to develop in front of the venturi tube and a suction at the throat of this tube. The excess pressure and the suction act on the underside and on the top of the diaphragm respectively. As soon as a certain minimum quantity of water is flowing through the venturi, the diaphragm is—by the combined action of pressure and suction—lifted to such an extent that the water-controlled gas valve is fully opened (against the restraining force of a spring). The gas thus flows to the burner and is lit by the pilot flame. The combustion gases rising from the burner flames heat the water as it flows through the heat exchanger. When the water is subsequently turned off (or if the supply accidentally fails), the pressure on the two sides of the diaphragm becomes equal, the diaphragm descends, and the spring-loaded water-controlled gas valve closes, so that the burner (except the pilot flame) goes out. If the gas is turned off (or if the pilot flame is extinguished by whatever cause), the bimetallic spring cools, and the ignition safety valve is closed by the spring, so that no unburned gas can escape even with the water turned full on.

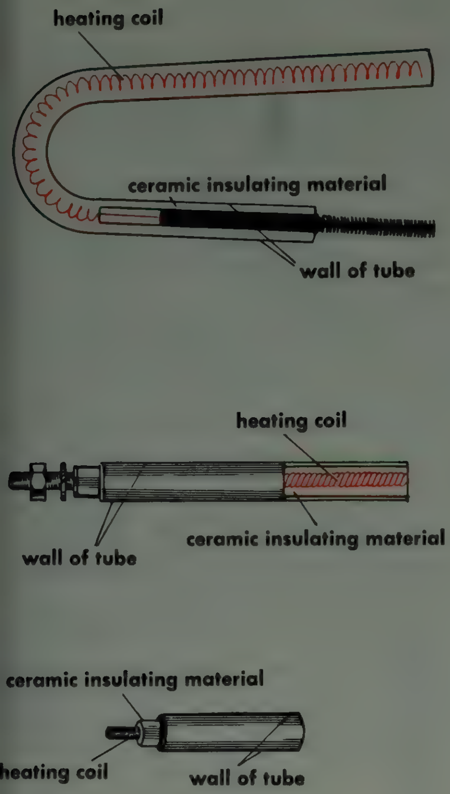

heat exchanger
pilot burner
burner
ignition safety valve system
gas flow controller
regulating screw (for adjustment if gas pressure is too low)
water-controlled gas valve
gas
gas control lever
venturi tube
diaphragm
pressure relief valve
bypass
water flow controller
temperature selector (closed)
water inlet
cold water outlet valve
hot water outlet valve

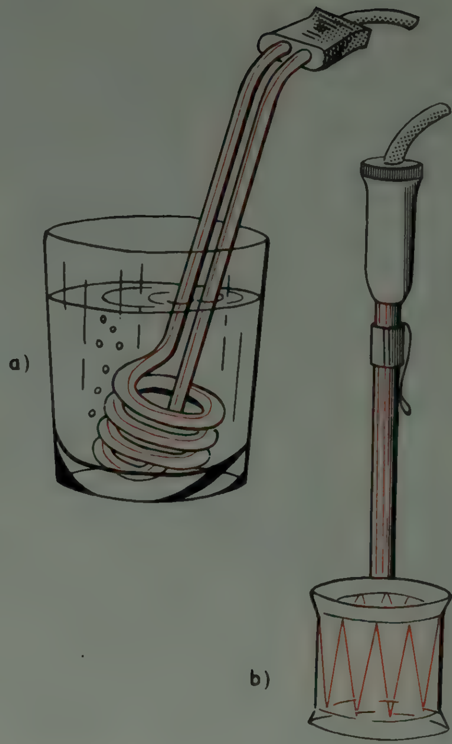
Thermal electric domestic appliances utilise the property of electric current of being able to develop heat when it encounters resistance. The electric current is conveyed by electrons. These transfer some of their kinetic energy to the atoms of which the conductor (the resistance wire) consists, causing these atoms to vibrate more violently about their respective equilibrium positions. It is this agitation of the atoms that manifests itself as a rise in temperature, i.e., heat. For a constant voltage, the heat produced by the current is proportional to the square of the current strength times the resistance of the wire, i.e., proportional to I^2R (Joule's law). To obtain a good heat output it is therefore necessary to use conductors having a high electric resistance. Various forms of tubular heating elements containing coiled resistance wire are illustrated in Fig. 1.

The simplest electric heating devices are *immersion heaters*, which may be either of the tubular type (Fig. 2a) or the annular type (Fig. 2b). The heat produced in the heating element is transferred to the surrounding liquid. The heat losses are small, and immersion heaters therefore have a high efficiency, nearly all the heat being given off to the liquid.

The principal component of an electric cooker is the *hot plate* (or boiling plate) (Fig. 3). Temperature control is usually effected by means of a bimetallic strip device (cf. page 26). A bimetallic strip consists of two strips of metal, with different coefficients of thermal expansion, bonded together. On being heated the two metals undergo different amounts of expansion, so that the strip curves and its free end (assuming the other end to be fixed) therefore deflects. This deflection, which is greater in proportion as the temperature is higher, can be used for actuating a contact or a switch. If the distance from the switch to the home ("cold") position of the end of the strip is made variable, the bimetallic device will be able to switch the current on and off over a whole range of operating conditions. Another device for utilising the difference in thermal expansion of two metals for performing a temperature-controlling function in domestic appliances is the Invar rod. Invar is an iron–nickel alloy with a very small coefficient of thermal expansion, so that it undergoes hardly any change in length on being heated. One end of such a rod is fixed in a slightly shorter brass tube. When the tube becomes hotter, it expands and increases its length, so that the end of the Invar rod disappears into the tube. When the tube cools, the end of the rod emerges again. The relative movement of the tube and the rod can be used for operating a switch (e.g., through suitable levers). A *water-boiling pan* (Fig. 4) may be equipped with a temperature-protective device (thermal release) which must automatically switch off the current in the event of overheating and then be capable of resetting to normal functioning of the appliance by pushing a button. a thermal release of this kind often comprises a bimetallic contact.

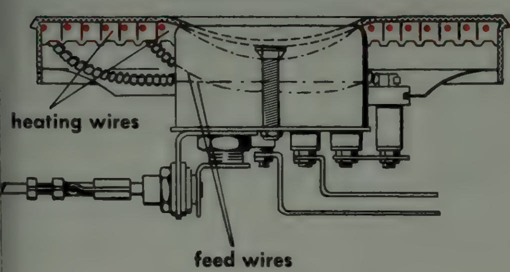
(Continued)





ı. 1 TUBULAR HEATING ELEMENTS USED IN SMALL DOMESTIC HEATING APPLIANCES

Fig. 2 IMMERSION HEATER: (a) *tubular type* (b) *annular type*



. 3 SECTION THROUGH A HOT PLATE

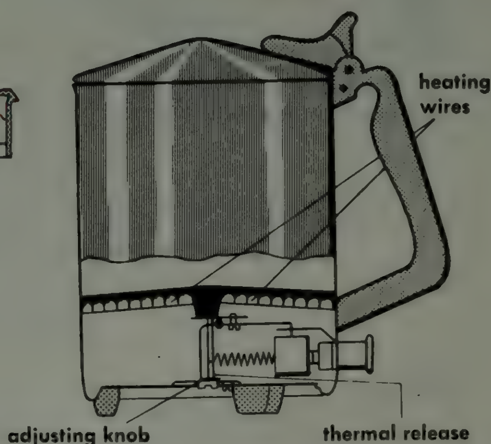


Fig. 4 SECTION THROUGH A WATER-BOILING PAN

With the immersion heater, water boiling pan and hot plate the transfer of heat is effected mainly by conduction and convection. In the *electric toaster* (Fig. 5), however, the heat is transferred by radiation. The slices of bread are inserted into two slots which have heating elements on both sides. When the toast is ready, a bimetallic device releases a spring, causing the slices of toast to "pop up".

In some warming pads the temperature can be pre-set to one of three maximum values, e.g., 80°, 70° or 60° C. When the pre-selected temperature is reached, the current is automatically switched off by means of a bimetallic device. When the temperature drops, the current is switched on again. The settings for these various temperatures are obtained by varying the degree of preheating of the bimetallic strip (Fig. 6). With the lowest preheat setting the highest temperature stage is obtained. In addition, a bimetallic safety device is provided, which breaks the circuit in the event of an excessively high temperature being reached (about 85° C).

Electric water heating appliances may be subdivided into storage heaters, boilers, and flow heaters (geysers). A *storage heater* comprises a water tank which is provided with efficient thermal insulation and whose contents are heated by a kind of immersion heater. Fig. 7a illustrates a small heater of this kind. The temperature control is "infinitely-variable" by means of a bimetallic controller of special design (Fig. 7b). When the bimetallic strip is heated, it curves upwards and thrusts against an actuating lever which is, at first, kept pressed down by a spring. When the force exerted by the bimetallic strip overcomes the counteraction of the spring, the actuating lever springs up and opens the contacts. With the heating element thus switched off, or when hot water is tapped from the tank (which is then replenished with cold water), the bimetallic strip cools and the operation is reversed, causing the contacts to be closed. By rotating the control knob for pre-selecting the temperature, the setting of the lever mechanism in relation to the bimetallic strip is varied, so that the device can be made to control the temperature over a fairly wide range (e.g., from 35° to 85° C).

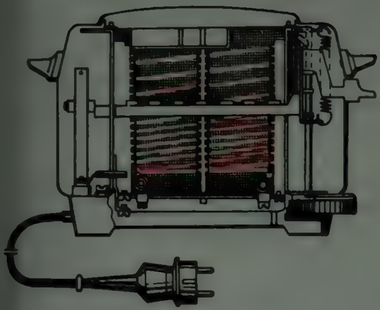


Fig. 5 ELECTRIC TOASTER

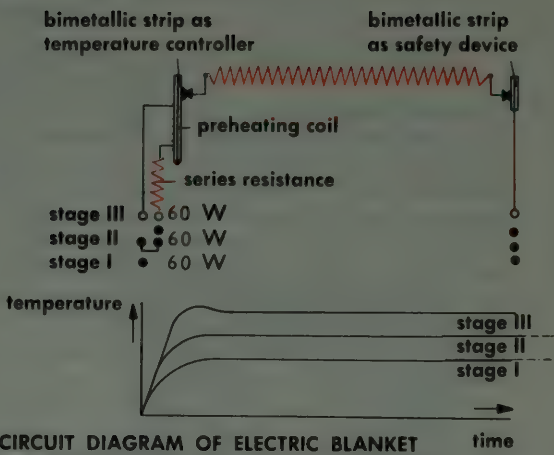


Fig. 6 CIRCUIT DIAGRAM OF ELECTRIC BLANKET WITH PRE-SELECTED TEMPERATURE CONTROL

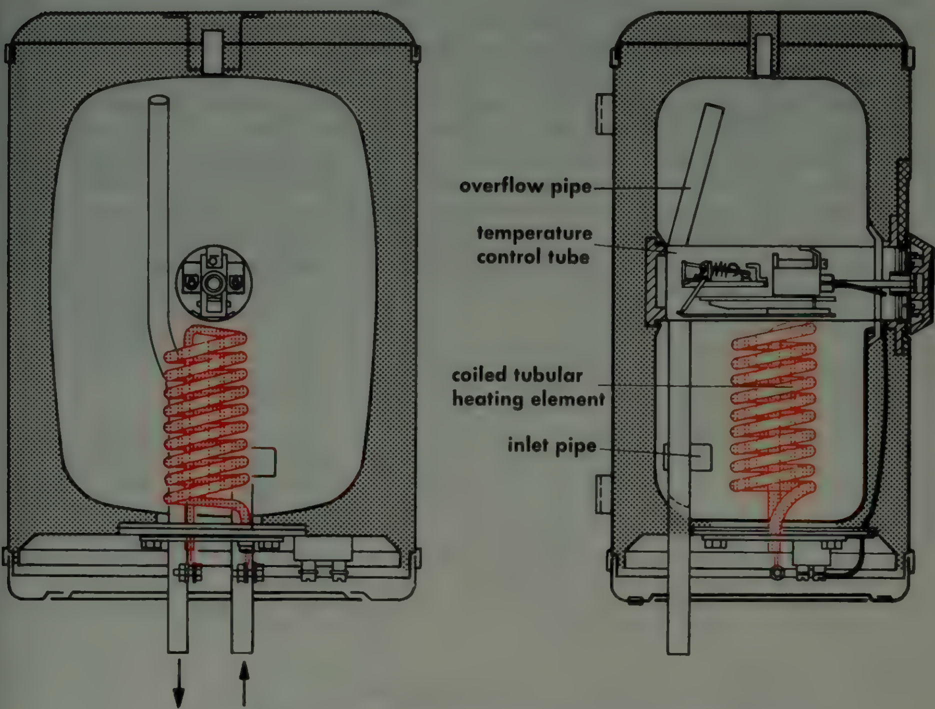


Fig. 7a SECTION THROUGH A SMALL STORAGE HEATER

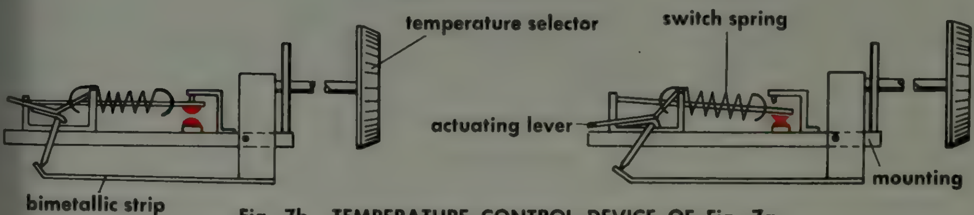


Fig. 7b TEMPERATURE CONTROL DEVICE OF Fig. 7a

A different temperature control device for hot-water storage heaters is illustrated in Fig. 8. It is installed in a protective tube beside the heating element and in close thermal communication with the hot water. Its principal component is an Invar rod (cf. page 304) which is fixed to an expansion tube. When it is heated, this tube expands, so that its end moves in relation to the end of the rod. This movement is transmitted through a lever mechanism to a mercury tube switch which tilts at a pre-set temperature and thus switches off the heating element.

In a *flow heater* (geyser) the water is not heated until just before it is actually needed. An appliance of this kind must heat the water very rapidly and must therefore have a very considerable heating capacity (about 12 kW, as compared with 2 kW for a small storage heater). See Fig. 9.

Electric *boilers* differ from storage heaters more particularly in having no thermal insulation. In this type of appliance, too, the water is heated only a short time before it is actually required. Once heated, it must be used quickly, for otherwise uneconomically large heat losses will occur. These appliances are relatively cheap and are moreparticularly suitable in cases where hot water is needed only at particular times.

A wide range of electric appliances for *room heating* purposes has also been developed. Wall-mounted (Fig. 11) and portable electric radiant heaters emit powerful heat rays which exercise their effect within a relatively small distance. The radiating unit comprises a parabolically curved reflector in whose line of focus a heating element is mounted. This reflector, more particularly if it is of the adjustable type, directs the heat rays in the desired direction. A heater of this kind does not efficiently heat up any significant volume of air. It has an effective "range" of only a few feet, and is used chiefly for auxiliary heating purposes (in bathrooms, bedrooms, etc.). More efficiently "space heating" is provided by the *fan heater* (Fig. 10). In this appliance the air drawn in by the fan is passed over a system of heating coils and then discharged into the room. A bimetallic cut-out is usually provided as a safeguard against overheating, more particularly if the fan fails to perform its proper function or if the free discharge of air is obstructed.

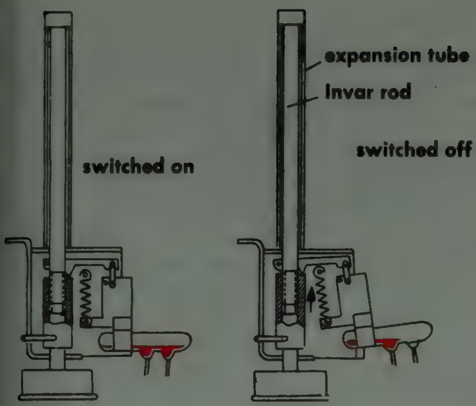


Fig. 8 TEMPERATURE CONTROLLER FOR HOT-WATER STORAGE HEATERS

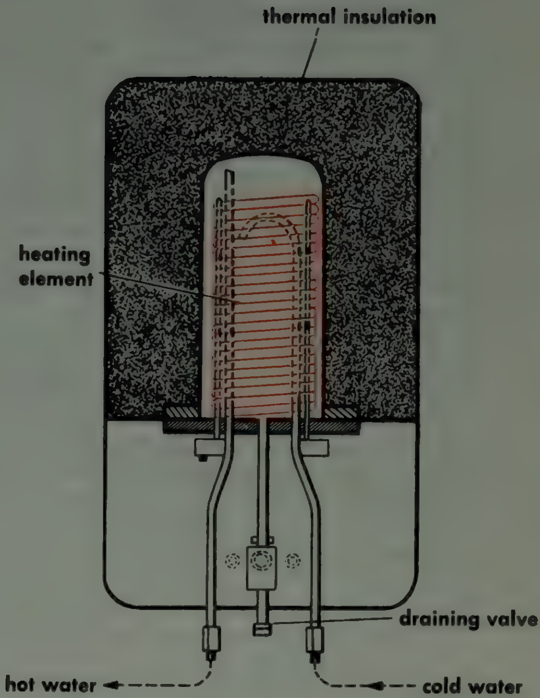


Fig. 9 FLOW HEATER

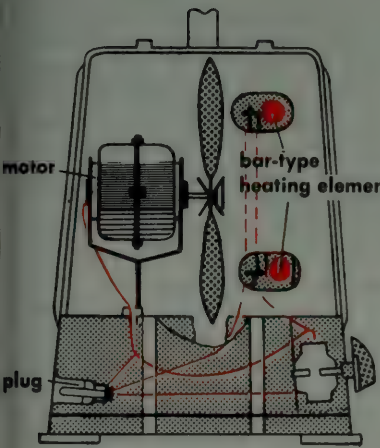


Fig. 10 FAN HEATER

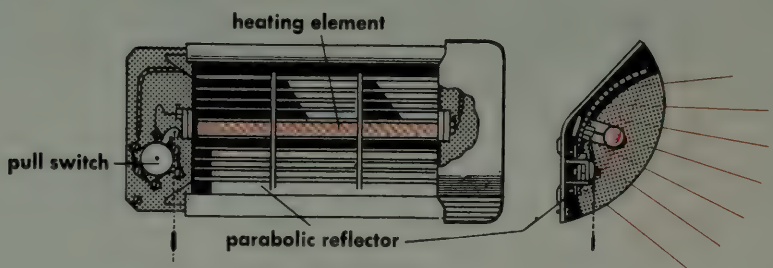


Fig. 11 RADIANT HEATER

Refrigerators lower the temperature inside them by extracting heat from the interior. Two laws of physics are utilised in achieving this: 1. The boiling point of a liquid, i.e., the temperature at which it is turned into vapour, depends on the ambient pressure. Thus at a pressure of 1 atmosphere (normal atmospheric pressure), water boils at 100° C, but at a pressure of 0.1 atmosphere it will boil at only 46° C. This means that, conversely, water vapour of, say 50° C and 0.1 atmosphere pressure can be condensed, i.e., converted back into liquid water, simply by increasing the pressure to, for example, 1 atmosphere. 2. On passing from the liquid to the vaporous state, every liquid absorbs heat and subsequently gives off this heat again on condensing. If we choose a liquid whose boiling point at normal atmospheric pressure is below the low temperature that we wish to obtain, such a liquid will already evaporate ("boil") at that low temperature and will absorb heat while it does so. This heat is extracted from the surroundings. Now if the vapour formed in this way is compressed, it will condense even at ordinary room temperature because the higher pressure produced by compression is associated with a higher boiling point. On condensing, it gives off heat. If the pressure is then reduced back to normal, the cycle can be repeated. In order to obtain the desired effect, so-called refrigerants are employed (these are liquids with low boiling points or liquefied gases, e.g., ammonia, ethyl chloride or Freon). Fig. 1 shows the cycle of operations: the heat needed for evaporation of the refrigerant is extracted from the refrigerating compartment, so that the temperature inside the latter goes down. Next, the refrigerant is condensed, giving off heat in the process, and then made to evaporate again.—There are two kinds of refrigerator:

Compression refrigerator (Figs. 2 and 3): The refrigerant, which is under low pressure, is evaporated in the evaporator. The latter is a coiled pipe installed in the freezer compartment. The evaporation lowers the temperature in the refrigerating compartment. A small compressor draws away the vapour, compresses it, and passes it to a condenser, where it parts with its heat. As a result of the combination of increased pressure and loss of heat, the refrigerant condenses. Finally, the now liquid refrigerant is expanded to the lower pressure and is returned to the evaporator. The temperature inside the refrigerator is regulated by a thermostat (see page 26) which switches the compressor motor on and off through a relay (see page 112).

Absorption refrigerator (Fig. 4): The absorption refrigerator operates without a compressor. The pressure is built up in a so-called boiler or generator, which is usually heated by electricity and is filled with water containing a high concentration of dissolved ammonia. When this solution is heated, the ammonia is driven off as vapour and the water remains behind. As this ammonia evaporation continues, the pressure rises until it is high enough to cause the ammonia vapour to condense in the condenser. Just as in the compression refrigerator, the liquid is then expanded by means of a special valve and thereupon evaporates again, absorbing heat (from the interior of the refrigerator) in doing so. The water which remains behind in the boiler after most of its ammonia has been driven out, and which is still very hot, is passed through a heat exchanger in which it gives off some of its heat. It then goes to the absorber, where it re-absorbs, and becomes saturated with, the pure ammonia vapour coming from the evaporator. The ammonia solution formed in this way is pumped back through the heat exchanger, where it absorbs heat from the hot water flowing from the boiler, to the boiler by a small circulation pump. The cycle of events then starts all over again.

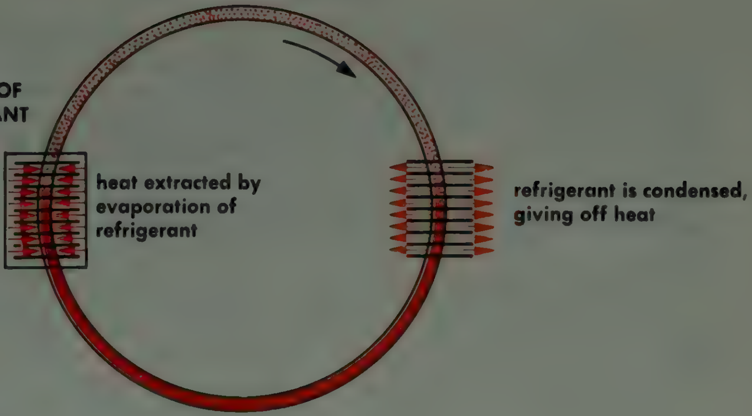


Fig. 1 CIRCULATION OF THE REFRIGERANT

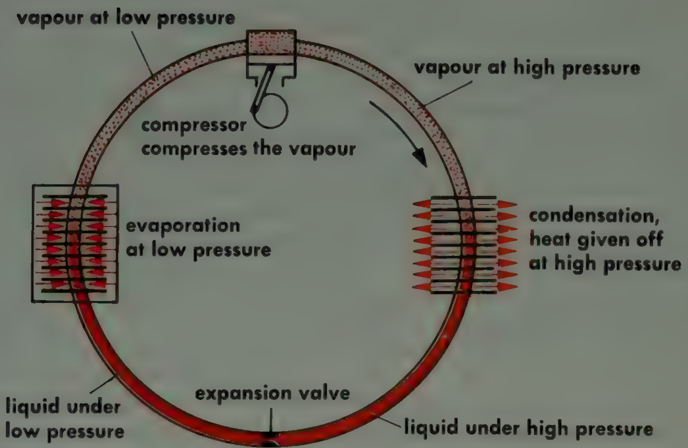


Fig. 2 OPERATING PRINCIPLE OF COMPRESSION REFRIGERATOR

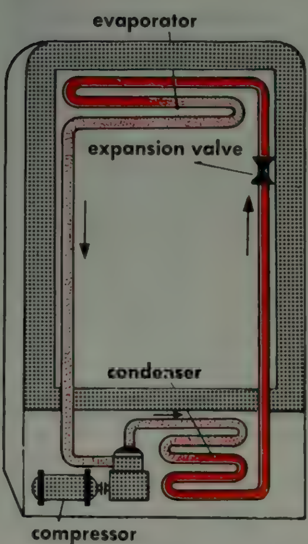


Fig. 3 COMPRESSION REFRIGERATOR

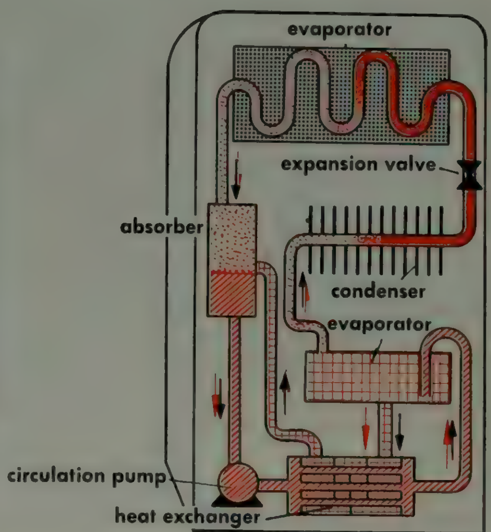
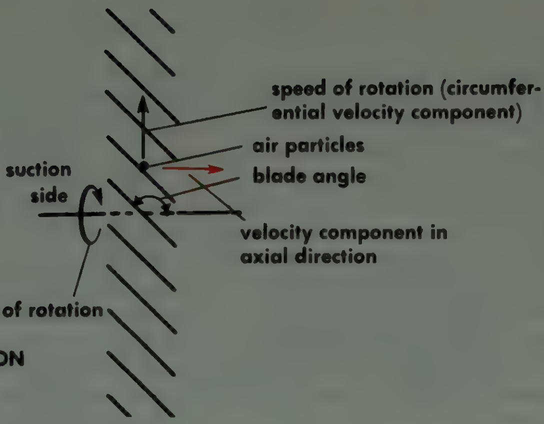


Fig. 4 ABSORPTION REFRIGERATOR

A modern vacuum cleaner develops its suction by means of a fan which discharges a powerful stream of air from the rear end of the casing. This sets up a powerful inflowing current of air which carries along any dust particles from the carpet or floor to which the suction nozzle is applied. The fan is usually driven by a small high-speed universal motor (i.e., a motor which can be worked either on alternating current or direct current). The fan has a large number of blades set at an angle. Their rotation sets up a flow of air in the axial direction (Fig. 1). The air stream is passed through a filter in which the dust is precipitated and collected without appreciably obstructing the flow.

Many vacuum cleaners have bag-type filters (Fig. 3) through which the air passes from the inside to the outside, the dust being collected in the bag, which has to be emptied from time to time. The disadvantage is that the suction power gradually diminishes because of the increasing air flow resistance as the bag fills up with dust. In the arrangement illustrated in Fig. 2 this disadvantage is obviated. Here the dust is precipitated in the dust collecting chamber in front of the filter. This system is used more particularly in good-quality hand vacuum cleaners; these have small dust collecting chambers, and it is therefore essential that they retain their suction efficiency unimpaired for as long as possible. In general, the power input stated on the rating plate of a vacuum cleaner does not necessarily provide a reliable indication of the suction performance. A sufficiently powerful airflow to carry along dust and grit particles must be set up. Such an air flow can only be induced by the suction developed by the fan. There are thus two factors involved: air flow rate (m3/min.) and suction (mm water column). These can be plotted against each other in a graph (Fig. 4), whereby a flow rate/pressure characteristic is obtained, which—depending on the type of fan—may be very steep or relatively flat. The suction performance is the product of these two factors. From the graph it appears that the suction is zero when the flow rate is maximum, and vice versa. At both these extreme points the suction performance is therefore zero. For a vacuum cleaner with a straight-line characteristic the best performance is obtained in the middle, i.e., at half the maximum suction and half the maximum air flow rate. In a well designed vacuum cleaner the various nozzle attachments must therefore be so dimensioned and shaped that the resulting performance is within the suitable working range.



g. 1 PRINCIPLE OF FAN ACTION

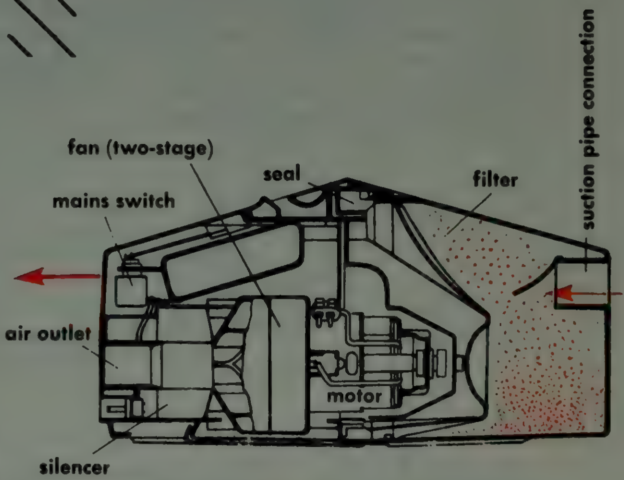


Fig. 2 VACUUM CLEANER WITH DUST COLLECTING CHAMBER

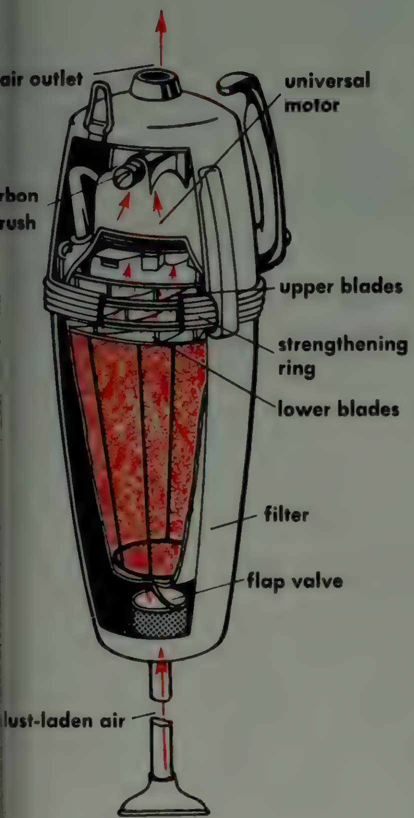


Fig. 3 VACUUM CLEANER WITH BAG FILTER

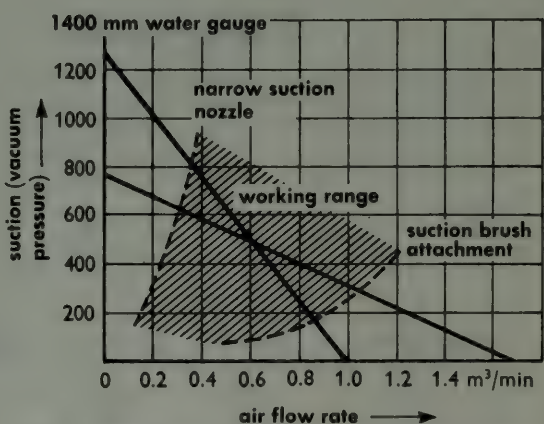
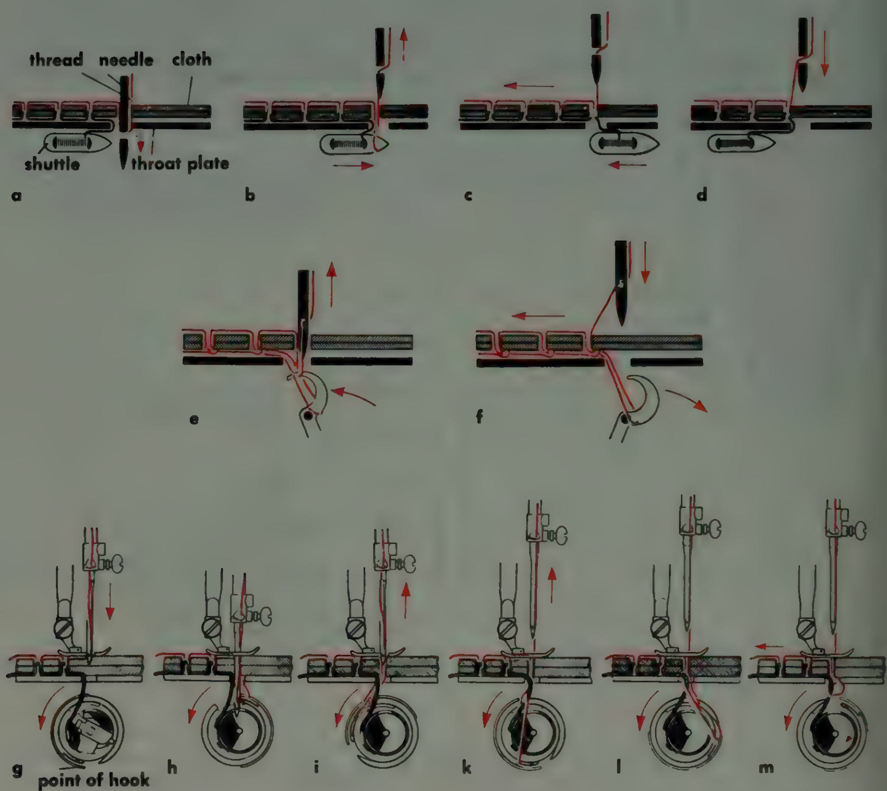
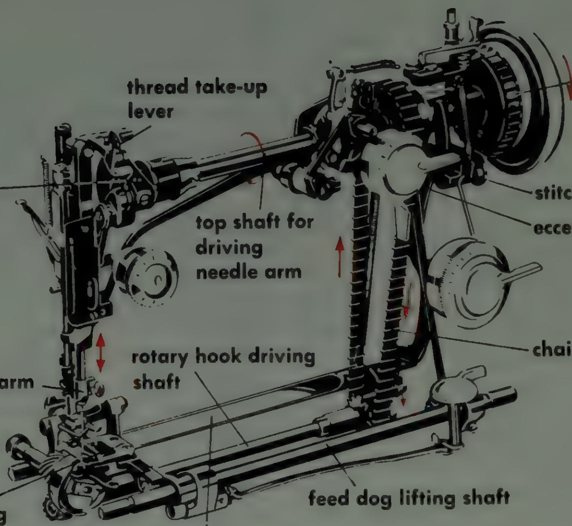


Fig. 4 AIR FLOW/SUCTION CHARACTERISTICS OF DIFFERENT VACUUM CLEANERS

The stitches made by a sewing machine are formed by two threads which are interlocked. In the vibrating shuttle machine the upper thread is carried by the needle, while the under thread is unreeled from the bobbin. The descending needle penetrates the fabric and carries the thread along (a). When the needle rises again, the thread forms a loop on the underside of the fabric. The shuttle, which contains the bobbin of under thread, goes through this loop and pulls the under thread along behind it (b). The shuttle thread is thus enclosed in the loop of the needle thread. The fabric is then moved forward ; while this is happening, the needle remains stationary and the shuttle returns to its initial position. This causes the slack loop to be pulled tight and close up, so that the two threads interlock in the middle of the fabric (c). When the forward movement of the fabric has ceased, the operation is repeated (d). This method produces the lock stitch, which forms a strong but rather rigid seam with no "give" in it.

A chain-stitch sewing machine produces seams having greater resilience. This machine works with only one thread, which is linked at the underside of the fabric by means of a gripper hook (e and f). A variant of this stitch is the overcasting stitch which enwraps the edge of the material.



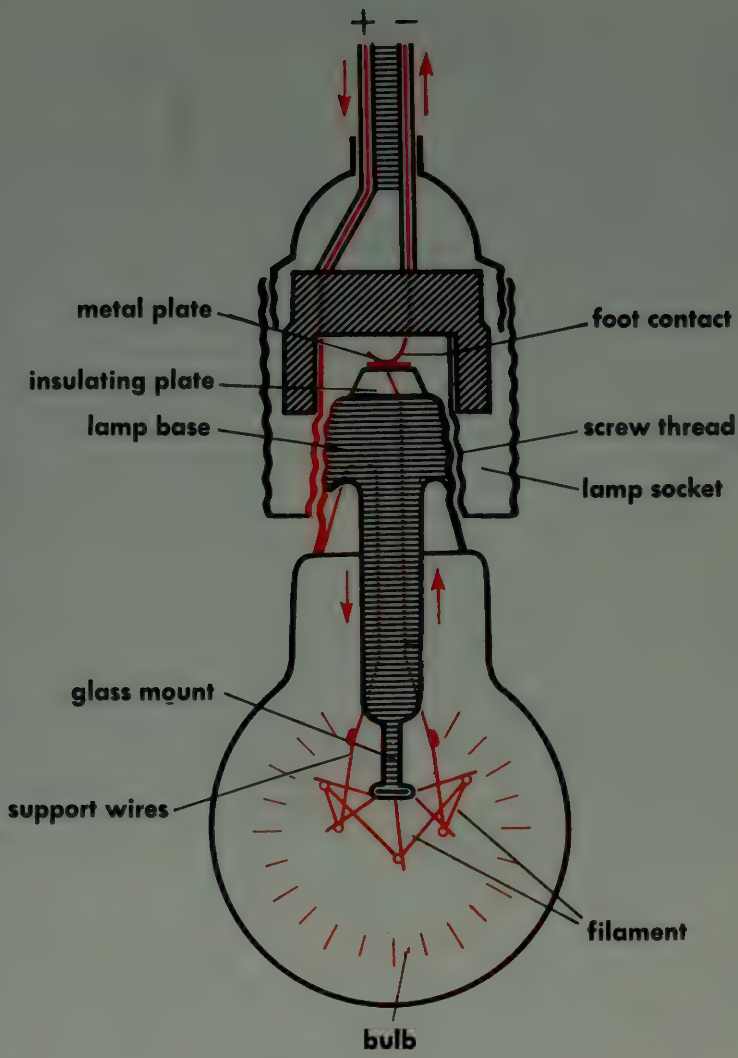


DRIVING MECHANISM OF A SEWING MACHINE

In modern domestic sewing machines the so-called rotary hook is frequently employed. This type of machine functions as follows. The needle descends through the fabric, and the point of the hook advances to meet the needle (g). The return movement forms the loop, and the point of the hook enters it (h). The hook enlarges the loop, the front of which is held in a recess in the bobbin case (i), while the hook pulls away the other side of the upper thread loop over the bobbin case (k). The loop slips off the point of the hook, while the thread take-up lever (see illustration below) pulls the excess thread up again (l). During the unwinding of the thread, the side of the loop that was held in the recess is released and the loop is pulled tight (m).

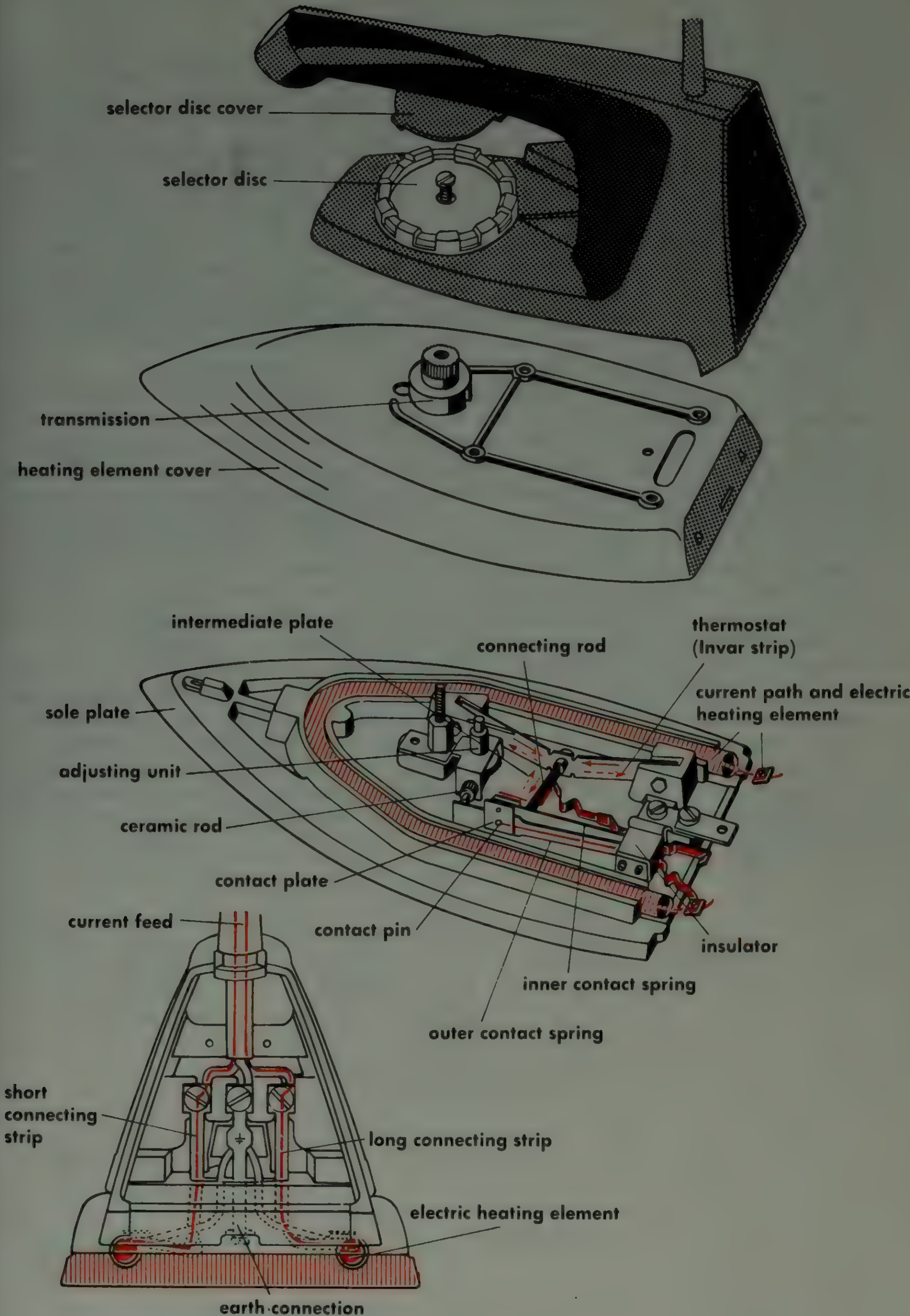
The present-day domestic sewing machine is usually driven by an electric motor through a chain drive. The rotary motion is transmitted through the top shaft to the crank drive for the thread take-up lever. In addition, the bottom shaft (rotary hook drive shaft) is driven by a chain from the top shaft. An eccentric cam mounted on the top shaft actuates eccentric rods and thus drives the feed mechanism under the base plate whereby the fabric is moved forward. The length of the stitch can be adjusted by means of the stitch regulator. The stitch is made longer or shorter by varying the eccentric stroke and thus varying the amount of rotation that the feed motion shaft undergoes at each stroke.

An incandescent lamp comprises an electrical conductor through which a current is passed which causes it to glow at white heat. The conductor is usually a wire, or filament, which is carried on a glass mount and whose ends are welded to thicker support wires (leads) through which the current is supplied to the filament. In order to prevent oxidation (burning away), of the filament by exposure to air, it is enclosed in a glass bulb, which is sealed together with the mount. The lead-in wires are sealed into the glass. The bulb is either evacuated, i.e., a vacuum is formed inside it, or it is filled with a neutral gas or gas mixture (e.g., nitrogen and argon). The filaments used in the early incandescent lamps were made of carbon. As it was not possible to raise the temperature of such filaments to white heat without seriously shortening the service life of the lamps, the light they gave was rather dim. For this reason carbon filaments were abandoned in favour of metal filaments. A suitable metal for the purpose was tungsten, which has a high melting point and can be heated to 3000° C. Tungsten is obtained, from its ore, in the form of a black powder, which is then processed, by sintering at about 1000° C in a neutral gas atmosphere, into pencil-size rods. The material is homogenised by hammering and stretched to rods about 18 in. long and $\frac{1}{8}$ in. in diameter. Further treatment is effected by drawing (p.44, vol.II). The drawing dies consist of pierced diamonds, by means of which it is possible to produce tungsten wire down to about $\frac{1}{100}$ mm (0.0004 in.) diameter. Some idea of the extreme thinness of such wire is conveyed by the fact that nearly 200 miles of $\frac{1}{100}$ mm wire can be produced from 1 lb. of tungsten. These extremely thin wires are then formed into doubly coiled (coiled-coil) filaments which are secured to the mount in the manner described above. The total length of the filament wire in a 15-watt lamp is about 0.75 m (30 in.); the first coil has about 3000 turns; this coiled wire is then coiled in 100 larger turns, so that the overall length of the filament is reduced to about 3 cm ($1\frac{1}{4}$ in.). The sealed bulb of the lamp is cemented into a metal base, which may be of various types, the commonest being the bayonet type and the screw type. The bayonet base is cylindrical in shape, with two small pins projecting from the sleeve to engage lock slots in the lamp socket; the two lead-in wires terminate at two metal contacts at the foot of the base. In the screw-type lamp (see illustration) one wire is connected to an insulated central metal contact plate at the foot, while the other is connected to the screw-threaded metal side of the base.


+
−
metal plate
foot contact
insulating plate
lamp base
screw thread
lamp socket
glass mount
support wires
filament
bulb

The modern electric iron comprises a sole plate, an intermediate plate, a cover with handle, an electric heating element, a selector disc, and a glow lamp. By means of the selector device the temperature can be set to values suitable for different fabrics (nylon, rayon, silk, wool, cotton, linen). This adjustment is achieved by means of a thermostat. The three-wire A.C. connection comprises an earthed safety wire. The heating element is shown in red in the accompanying diagram. The current goes through the element, the outer contact spring and contact pin, the inner contact spring and contact pin, the connecting rod, and the flexible metal strip.

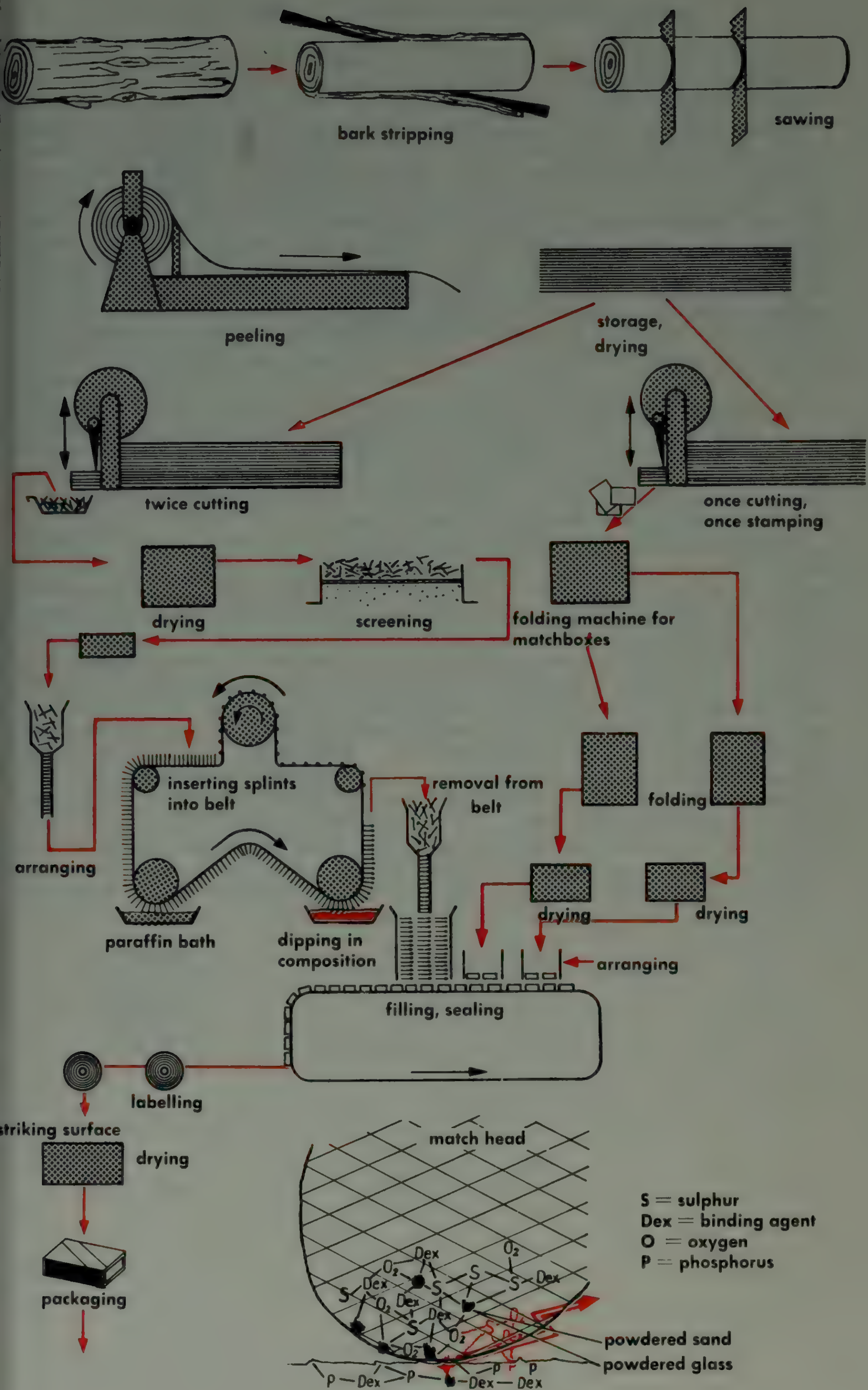
When the selector dial is, for example, set to "linen", the glow lamp lights up, and the current heats the heating element to the required temperature. The lamp then goes out. The iron is thereafter kept at constant temperature by the thermostat. For example, when the iron cools a little, the temperature of the freely movable intermediate plate, to which the thermostat is attached, also goes down. As a result of this, the intermediate plate contracts and causes the thermostat strip, which is made of Invar (a nickel-iron alloy with a very small coefficient of thermal expansion, so that it hardly expands or contracts due to temperature changes), to buckle outwards and push the connecting rod against the inner contact spring thrusts the latter against the outer contact spring, so that the circuit is restored and the heating element switched on again. The glow lamp also lights up. When the pre-set ironing temperature has been reached again, the thermostat causes the contacts to separate and thus break the circuit. The glow lamp is thereby also switched off. The switching on and off of the lamp indicates that the thermostat is functioning and automatically keeping the temperature constant. Depending on the setting of the selector dial, the heating element is switched on and off at a higher or lower temperature. Rotation of the disc moves the rod of ceramic material in the axial direction. The outer contact spring is so installed that it tends always against the inner contact spring. When the iron heats up, and the Invar strip consequently elongates, the outer spring follows the inner spring until its end encounters the ceramic rod. This causes the contacts to separate. The farther forward this rod protrudes, the sooner this occurs, i.e., the lower the temperature at which the heating element is switched on and off. A thermostat of this kind controls the temperature to within about 10° C accuracy.


selector disc cover
selector disc
transmission
heating element cover
intermediate plate
connecting rod
thermostat (Invar strip)
current path and electric heating element
sole plate
adjusting unit
ceramic rod
contact plate
contact pin
insulator
inner contact spring
outer contact spring
current feed
short connecting strip
long connecting strip
electric heating element
earth connection

In the manufacture of safety matches, softwood logs (e.g., poplar) are peeled into a thin continuous shaving, or veneer, about $\frac{1}{10}$ inch thick (veneer process). The ribbon of wood is then cut up into splints at a rate of about two million per hour. These splints are soaked in a bath of sodium silicate, amonium phosphate or sodium phosphate and then dried. This impregnation prevents afterglow. Next, the splints are fed into a continuous match machine in which the splints, standing on end in wide belts, are passed through a paraffin bath, which treatment aids ignition. Next the machine dips the ends of the matches in a liquid composition which becomes the striking head when dry. This composition consists of the oxygen carrier (potassium chlorate, lead oxide, potassium chromate, manganese dioxide, etc.), the inflammable ingredient (sulphur, etc.), frictional additives (powdered glass), colouring matter, and binding agents (dextrin, gums). The striking surface on the matchbox consists of powdered glass, red phosphorus, colouring matter, and binding agents.

Safety matches are so called because they can be ignited only by friction against the striking surface of the box. On the other hand, there are universal matches which have heads of such composition that they can be lit by striking them on any friction surface. The heads of such matches have a somewhat different chemical composition from ordinary safety matches. The round wooden matches manufactured in America have two-colour heads and are produced by a double dipping process. The large bulb of the match consists of mainly inert substances, while the "eye" contains the readily ignitible ingredients. The bulb having the larger diameter prevents the sensitive "eyes" from rubbing together after packing. Another American invention, now in worldwide use, is the paper book match. These matches are packed in printed cardboard folders with a striking surface on the outside.

When the head of a safety match is rubbed against the striking surface, frictional heat is generated at a small area of the head. The heat liberates oxygen from the oxygen carrier ingredient, and this oxygen combines with the sulphur (sulphur dioxide is formed in this reaction), whereby additional heat is evolved, causing more oxygen to be liberated and react with sulphur. The chemical process thus initiated by friction proceeds very rapidly, so that the entire match is soon alight. The paraffin-impregnated splint also catches fire. The impregnation of the splint prevents it from continuing to glow after the flame is extinguished.


bark stripping
sawing
peeling
storage,
drying
twice cutting
once cutting,
once stamping
drying
screening
folding machine for
matchboxes
inserting splints
into belt
removal from
belt
folding
arranging
paraffin bath
dipping in
composition
drying
drying
arranging
filling, sealing
labelling
striking surface
drying
packaging
match head
S = sulphur
Dex = binding agent
O = oxygen
P = phosphorus
powdered sand
powdered glass

The first practical typewriter by present-day standards was commercially produced in America in 1874. This was the Remington, based on the machine invented by Sholes. The mechanical typewriter generally functions on the following principle: when any particular key of the keyboard is pressed, a system of levers and linkages (Fig. 1) causes the corresponding type bar to strike the paper through an inked ribbon. The imprint of the raised type on the type bar is thereby formed on the paper which is partially wrapped round the cylinder (or platen). The type bars all strike the paper at a common centre. Each time a key is released after being pressed, the carriage on which the cylinder is mounted moves a certain fixed distance to the left corresponding to a width of a letter. This motion is arrested by an escapement mechanism (pawl and ratchet wheel) which is briefly released each time the carriage moves one step to the left. When the end of the line is reached, the carriage is returned to the right by the typist. This is normally done by pusing a lever which works the line-spacing mechanism, i.e., it rotates the cylinder a certain fixed amount. Each type bar carries two types, e.g., a capital and a small letter. The capitals are typed by pressing a special key which operates the cylinder-shift mechanism whereby the capital instead of the small letter strikes the paper when the key corresponding to any particular letter is pressed. A typewriter embodying this now universal principle is known as a shift-key typewriter. The first shift-key typewriter was produced in 1878. Its early rivals were so-called single-key machines; they had twice the number of keys—one for every character, whether capital or small letter.

An important advance was the development of the electric typewriter, which is basically a mechanical typewriter in which the typing stroke is powered by an electric motor drive (Fig. 2). The key stroke, the carriage motion and other controls are initiated by touching the proper key. The motor rotates a drive roller at constant speed. When the typist presses the key a short distance, a cam (usually made of nylon) is brought into contact with this roller. The latter moves the cam along with it by friction. This causes the cam lever to move back and thereby actuate the type bar. All the functions of the electric typewriter are performed on this or a similar principle, whereby a 95% saving in physical energy expenditure is effected, so that operator fatigue is greatly reduced. Another advantage of the electric typewriter is that it makes for faster and more uniform typing. All the strokes are applied with equal power, since they are independent of the amount of force actually applied by the typist. Despite various refinements and improvements of detail, the basic functioning principle of the typewriter remained unchanged for many years until the advent of the typing-head system (see page 324). The so-called "noiseless" typewriter is a mechanical typewriter utilising a special type bar to reduce the noise of the impact of the type bar on the paper. It is not, of course, absolutely noiseless.

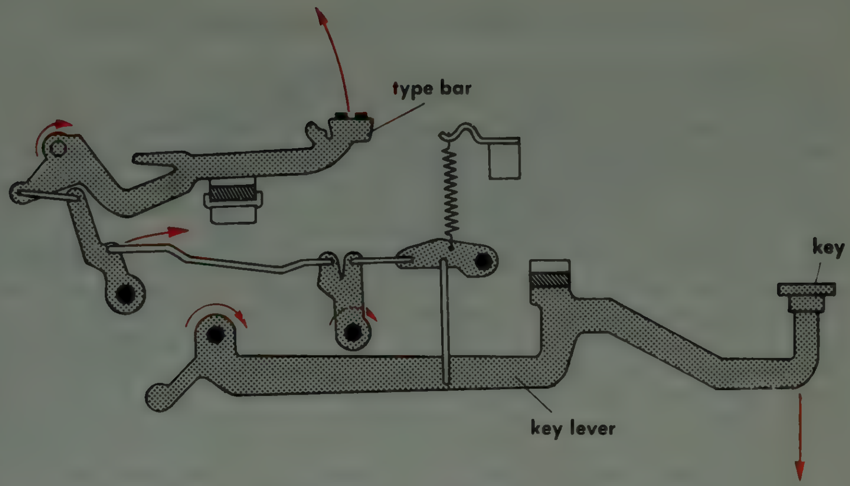


Fig. 1 MECHANISM OF A MECHANICAL TYPEWRITER

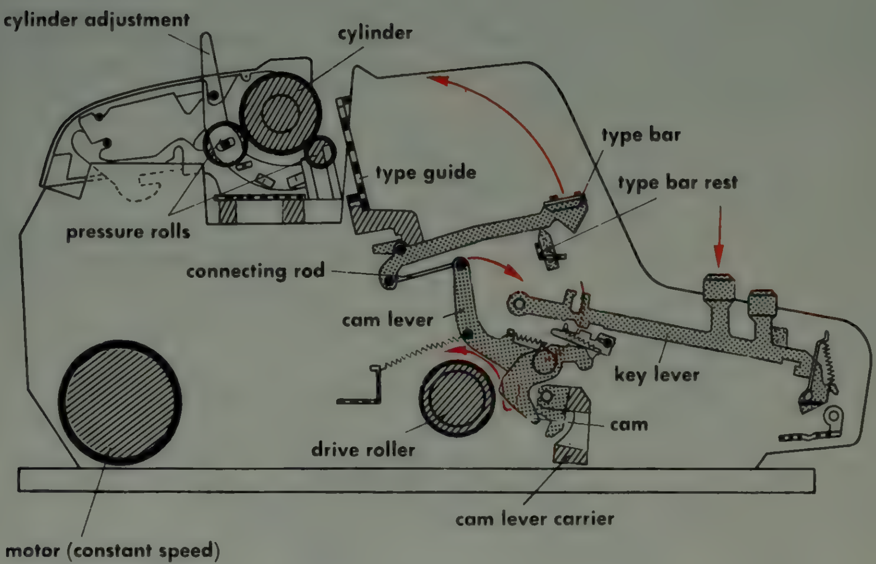


Fig. 2 SECTION THROUGH AN ELECTRIC TYPEWRITER

Instead of having 44 type bars, a machine of this kind has a spherical typing head of about $1\frac{3}{8}$ in. diameter, provided with 88 characters and moves along the line as it types. The head is made of nickel-plated moulded plastic. Owing to its light weight, combined with toughness and resilience, it can perform lightning-quick movements. It can speedily be replaced by another head whenever a different type of lettering is required for a particular purpose. The 88 characters are arranged in four rows round the head. In Fig. 1 the type marked by a cross is positioned ready to strike the paper on the cylinder. The actual striking is performed by a simple mechanism. To type a character situated at another of the 88 positions on the typing head, the latter must be appropriately swivelled and tilted so as to bring this other type into the striking position. These tilting and swivelling movements are performed by means of two steel band systems which begin and end at the slide which carries the typing head and inked ribbon holder. The tilting mechanism provides for four different positions; the swivelling mechanism has to move the head to 22 positions (11 with small letters on the front; 11 with capital letters on the back). To type any particular character, the machine must therefore select the appropriate position from four tilting and 22 swivelling positions. This is done by means of a selector system (Fig. 3). Under each key is a selector lever with projections on its underside. For each key these projections are arranged in a different combination. On being depressed by the key, the selector lever is moved forward by a rotating shaft of special shape. The projections on the underside thrust against various selector rods which operate through a system of selector catches and thereby bring about the requisite movements of the swivel arms of the steel band systems. As soon as the selected character is in the striking position facing the cylinder, the typing head is momentarily locked so as to prevent tilting and swivelling. The stroke is then delivered.

In this way a typing speed of $15\frac{1}{2}$ strokes per second, i.e., 930 per minute, can be attained. A notable feature is that the proper functioning of the machine cannot be upset by an excessively rapid succession of typing strokes by the typist. If a stroke is followed too quickly by another stroke, the latter is "stored" until the previous stroke has been duly completed.

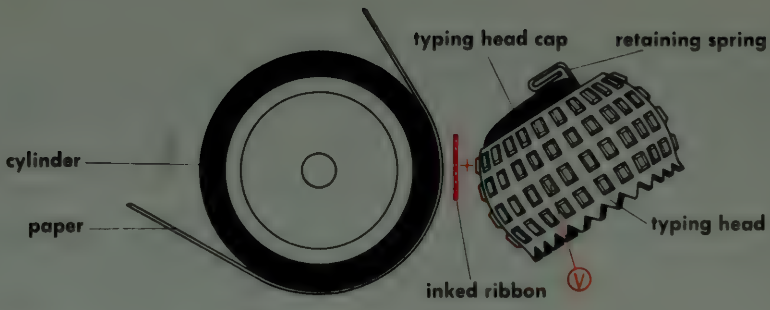


Fig. 1 TYPING HEAD AND CYLINDER (*side view*)

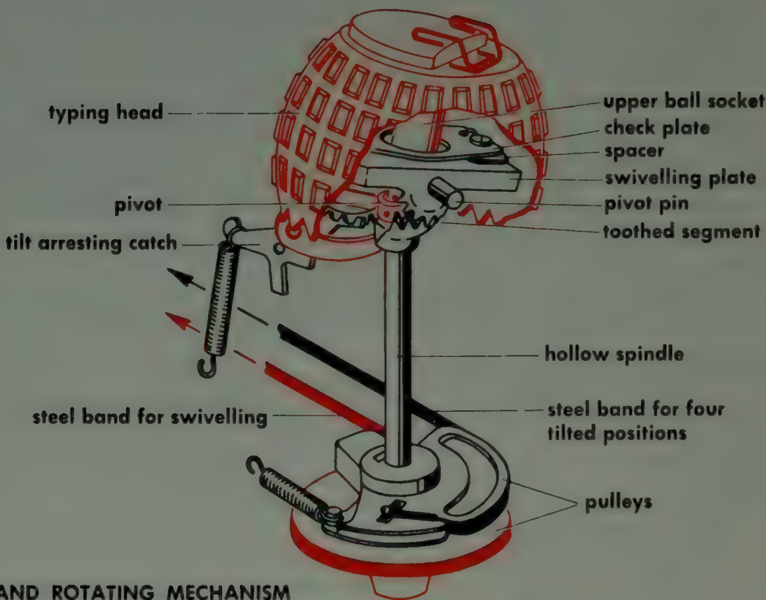


Fig. 2 TILTING AND ROTATING MECHANISM

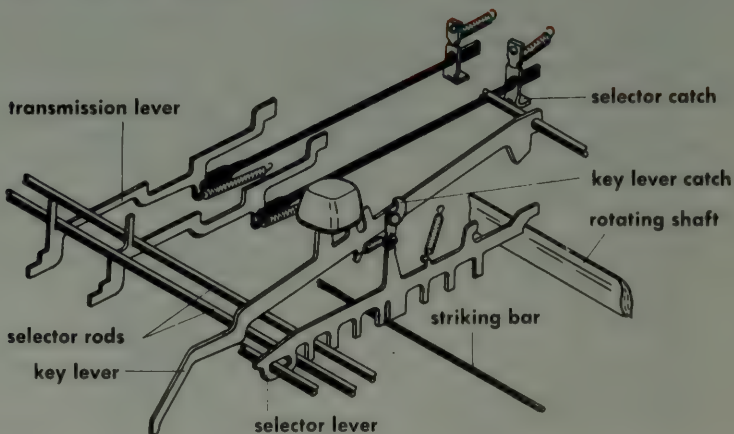


Fig. 3 SELECTOR SYSTEM

The primary function of a cash register is to record cash transactions on a strip of paper and add up the figures. For this purpose the machine is provided with a setting mechanism from which the figures are transferred to an adding mechanism, an indicator device, and a printing device. The latter prints a check strip, which remains in the machine, and a receipt slip, which is handed to the customer. A subsidiary function of a cash register is to serve as a till in which the money is kept. For this purpose it is provided with one or more cash drawers.

There are, in the main, two types of cash register which differ from each other in the setting mechanism and, consequently, in the entire adding and recording mechanism: (1) cash registers with lever action, in which the amounts are fed into the machine by setting various levers; (2) cash registers with key action, in which this is done by pushing various keys.

Lever-type cash register (Fig. 1):

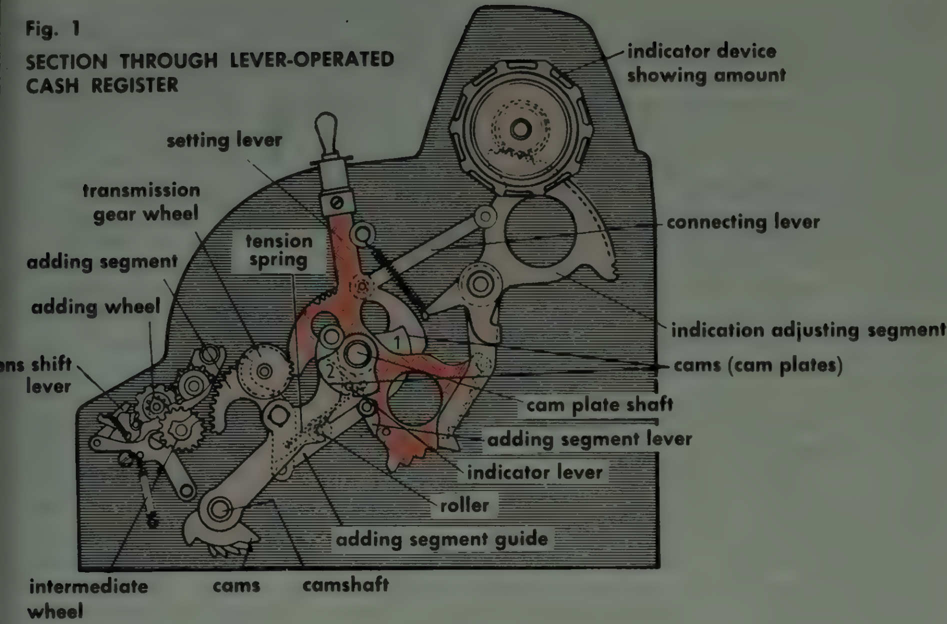
When the lever for setting the amount is shifted to the desired position, the amount appears in a window on the back of the machine. At the same time, a cog-wheel rotated by the setting lever sets the printing mechanism to the selected amount. In addition, the setting lever, acting through a tension spring, shifts the adding segment lever and the adding segment guide to the working position. By pressing a button, an electric motor is started, which drives a camshaft in the clockwise direction. Cam 1 causes the adding segment to engage with the intermediate cog-wheel of the adding mechanism; cam 2 and the roller transfer the amount to the adding wheel. The camshaft is a frequently used device for producing reciprocating motions of various kinds. depending on the shape of the cams. The principle of the cam action is illustrated in Fig. 1a. It is also used for working the valves in internal combustion engines (cf. p. 188, vol. II). With regard to the adding mechanism see also page 328.

Key-type cash register (Fig. 2):

When a key is depressed, the electric motor is started which drives the camshaft anti-clockwise. This causes the block, which is rotatably mounted on the shaft, to thrust against the sickle-shaped lever, so that this lever rotates clockwise about its pivot. This causes the swivel arm to swivel until it is stopped by the key bar; as a result of this, the bottom pivot of the sickle-shaped lever becomes a fixed point. The sickle-shaped lever is thus compelled to rotate in the clockwise direction. In doing this, it carries the counting arm along with it until this arm, too, is arrested by the key bar. Acting through the link rod 1, the counting arm moves the adding segment into position. The pin and segment 1 rotate the type transfer wheel and thereby set the printing wheels in position. The link rod 2, attached to the segment 1, works the indicator roll through the agency of the segment 2. The adding segment engages with the intermediate cog-wheel of the adding mechanism. Then the sickle-lever differential opens, and the counting arm and the swivel arm return to their initial positions. The amount is transferred both to the top adding wheel and to the bottom sectorised adding wheel. On release of the adding segment, the camshaft rotates and performs the transfer of the tens (cf. page 328).

Fig. 1

SECTION THROUGH LEVER-OPERATED CASH REGISTER



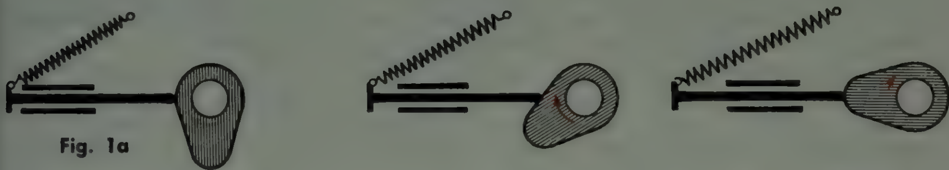
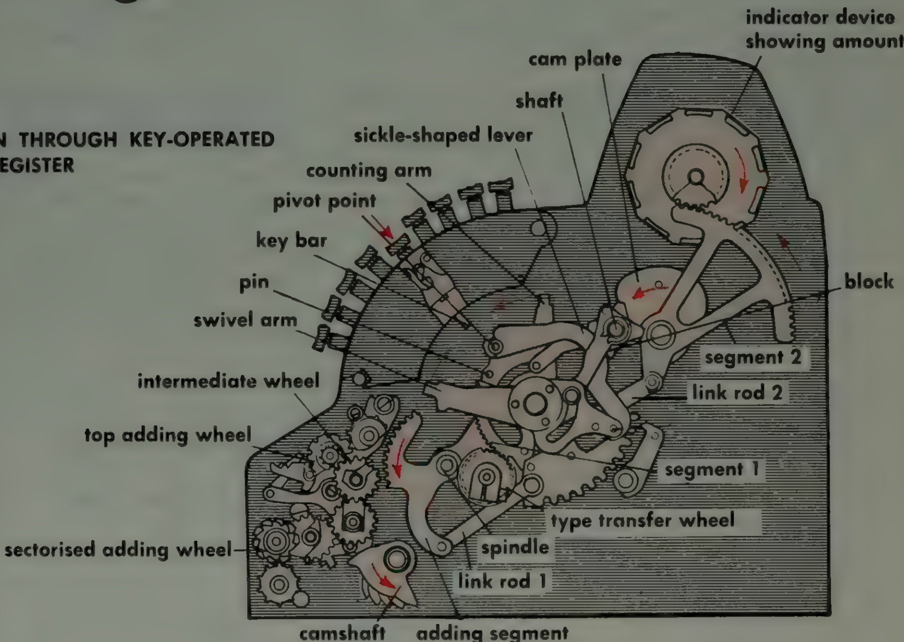


Fig. 1a

Fig. 2

SECTION THROUGH KEY-OPERATED CASH REGISTER



A desk calculating machine for performing addition and subtraction operations comprises the cog-wheel mechanism and the product register which records the result of the calculation. The product register is a counter mechanism with "tens" transfer function. Multiplication is carried out by continued addition, subtraction by reversal of the direction of rotation, and division by continued subtraction. A number is fed into the mechanism by pressing keys on a keyboard and is added to, or subtracted from, the number already in the product register. The mechanism is illustrated in Fig. 1. Under the keyboard are ten racks (toothed rods), each of which has on its underside a pin which engages with a groove in a cross-bar, the "proportional lever". When the machine is set to "addition", the rearmost rack (as shown in Fig. 1) is immovably locked. When the handle is turned through one revolution, the proportional lever is swivelled around the pin of this rack to the right and back again. In the course of this operation the front rack moves nine places, the next rack immediately behind it moves eight places, etc. Over the racks are mounted as many square shafts with slidable cog-wheels as there are places in the product register. Associated with each two keys of the keyboard is a slidable cog-wheel on the square shaft. This cog-wheel is moved by a sliding mechanism. In the neutral position only the locked rack is meshed with a cog-wheel. If, for example, the key bearing the figure 3 is pressed, this cog-wheel is disengaged, and another cog-wheel meshes with the rack which, on actuation of the handle by the machine operator, moves a distance corresponding to three teeth (Fig. 2). At the start of the forward motion, the coupling between the product register and the square shafts is moved into position. As a result of this, in each decimal place the counting wheel is rotated a number of teeth which corresponds to the figure on the keyboard. If the number exceeds 9 the counting wheel moves a marker slide for the tens transfer operation (see below). The coupling is disconnected before the racks reverse their direction of rotation.

When subtraction is carried out, the front rack is immovably locked. Now each rack moves to and fro a distance corresponding to (9-n) teeth instead of n teeth as in the case of addition. Instead of actually subtracting a from b, i.e., $a - b$, the machine performs the operation $a + (999 \ldots 999 - b)$. At the end of the subtraction the result is increased by one unit. The result $a + (1000 \ldots 000 - b)$ is equivalent to $a - b$ on the machine as the addition of $1000 \ldots 000$ affects only the next higher place, not present on the machine.

In the second half of the revolution performed by the operating handle the "tens" transfer is effected. Eccentric cams mounted in a staggered arrangement on a shaft successively raise the "tens" levers in the product register, starting with the units place. When the marker slide has been shifted, the "tens" lever can no longer slide between the cog-wheels; instead, it deflects to the left and thereby moves the cog-wheel of the next higher position a distance of one tooth. The displacements necessary for effecting a "tens" transfer are indicated by dotted red arrows in Fig. 4.

As already stated, multiplication and division are carried out as continued addition and subtraction respectively. In a multiplication the multiplier is registered in a special register called the revolution register. The multiplicand is set on the keyboard. The revolution and product registers can be moved sideways on a carriage to facilitate calculation with multi-digit multipliers. When the product register moves n places, one revolution of the operating handle adds 10^n times the number set on the keyboard. In that case the corresponding place of the revolution register must be reduced by 1. Fully-automatic machines perform the necessary displacements and terminate the calculation when the revolution register is at zero. Division is performed on similar lines.

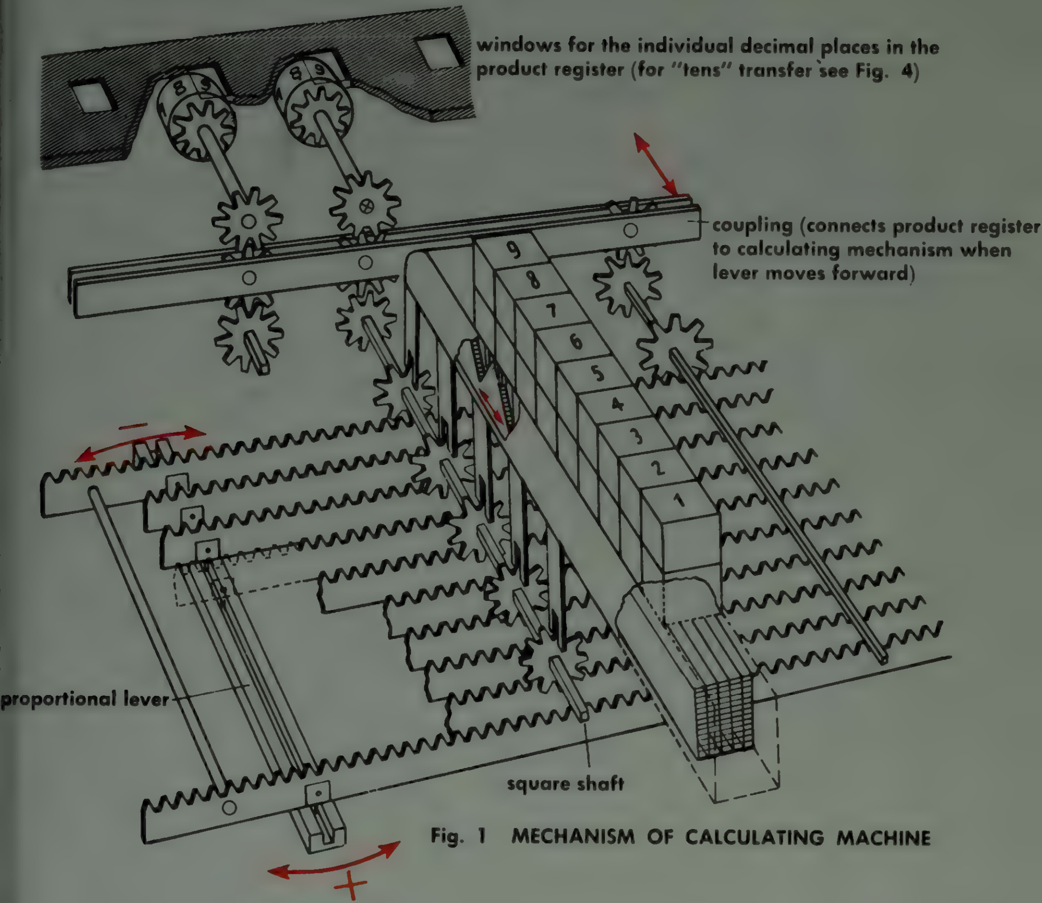


Fig. 1 MECHANISM OF CALCULATING MACHINE

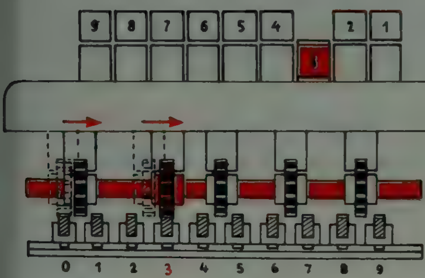


Fig. 2 KEY "3" DEPRESSED

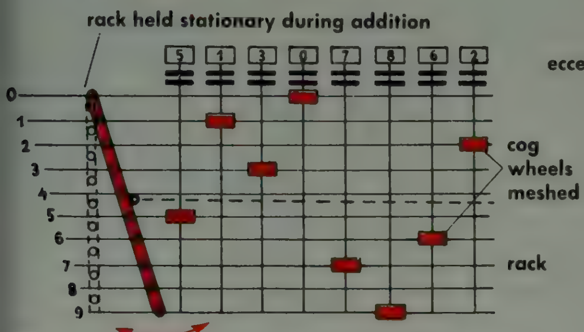


Fig. 3 MOVEMENTS OF LEVER AND RACKS
(*cog wheels not in mesh are not shown*)

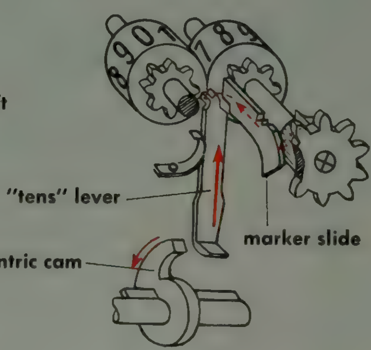


Fig. 4 "TENS" TRANSFER IN PRODUCT REGISTER

Print-out drums in a cash register, Injecta Teufenthal, Switzerland
Photo Roland Schneider, Len Sirman Press

Punched cards (like punched tape) can be used for effecting the input and output of data in mechanised data processing. Such cards can be sorted at rates of over 100,000 cards per hour and arranged in various ways.

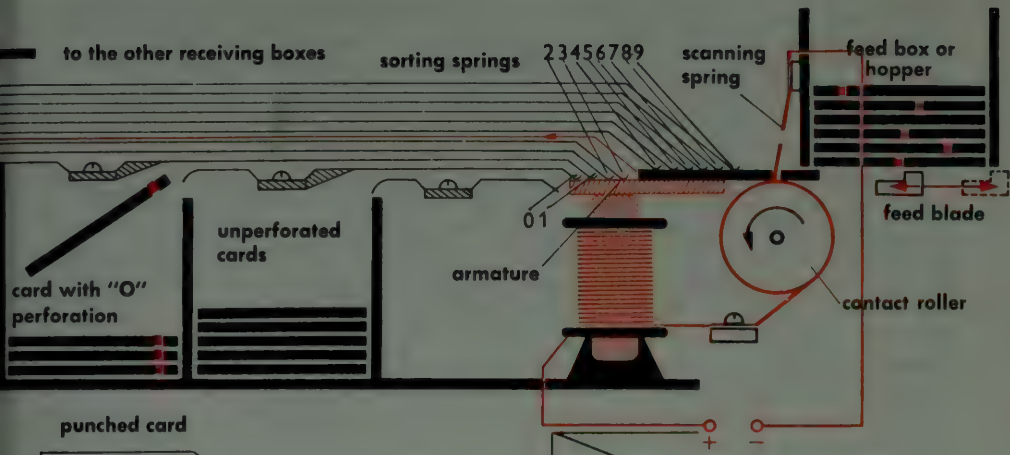
A punched card is subdivided into columns whose rows contain the figures from 0 to 9. For example, if the number 403 has to be recorded in the first three columns of a card, the row 4 in the first column, the row 0 in the second column, and the row 3 in the third column must be punched. The usual punched cards comprise 80 columns. In addition, appropriate holes for conveying non-numerical code information can be punched in rows in which the places are not numbered.

When the punched cards are passed through any of various machines, the punched holes cause transmission of impulses. Basic machines are the punching machine which punches data into the cards, the sorting machine which sorts out cards according to various classifications, and the tabulating machine which prepares printed reports from the sorted cards.

Figs. 1 and 2 show the operating principle of a sorting machine. The stack of cards for sorting is placed in the feed hopper on the right, in which to-and-fro-moving feed blades seize the bottom card and push it under the transport roller (Fig. 2). The card passes between a contact roller and a scanning spring. The latter scans the rows 9 to 0 in the selected columns. In the neutral position the sorting springs are close above the card transport track. Cards which contain no punched hole in the scanned column will pass unhindered under all the sorting springs and fall into the receiving box for unperforated cards. On the other hand, if the column contains a hole, the card concerned will slide along under the sorting springs only until the hole reaches the scanning spring. When that happens, an electrical circuit is completed which energises an electromagnet, whereby an armature, on which the sorting springs are resting, is pulled down. Those sorting springs whose front edges have then not yet been reached by the card will thereby drop below the level of the card, with the result that the card slides in between two of these springs and is thus delivered to the appropriate receiving box.

A stack of cards may also be sorted with regard to the value of a multi-digit number. This involves several successive sorting operations (Fig. 3). First, the cards are sorted according to the figure in the units column. Before the second sorting operation takes place, the cards must be stacked in the feed box in the sequence of their end digits (first the end digit 0, then 1, etc. up to 9). When the cards are now sorted with regard to the "tens" digit, it is thus assured that the cards with the lowest end digits are lowest down in the relevant "tens" box; the other cards are stacked upon these in the sequence of their respective end digits. In the next sorting operation the cards are sorted with regard to the "hundreds", then the "thousands", etc. Thus, with only four sorting operations 10,000 cards have been arranged in their correct order.

Fig. 1



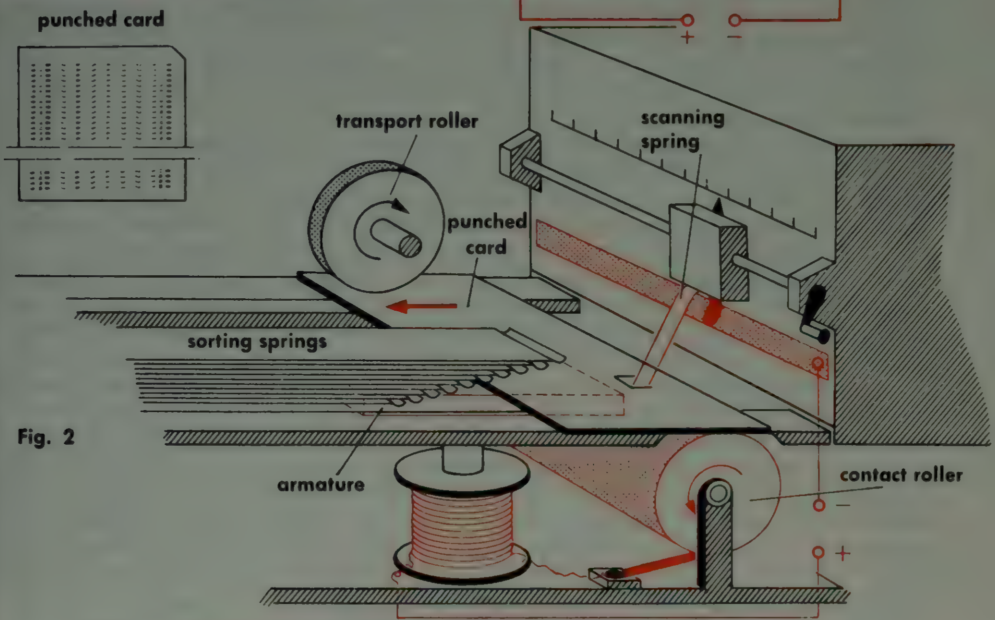
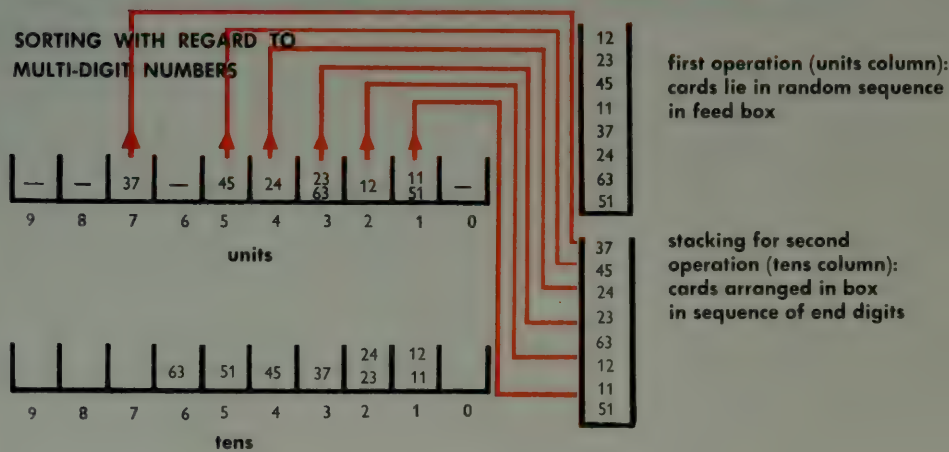


Fig. 2

Fig. 3 SORTING WITH REGARD TO MULTI-DIGIT NUMBERS

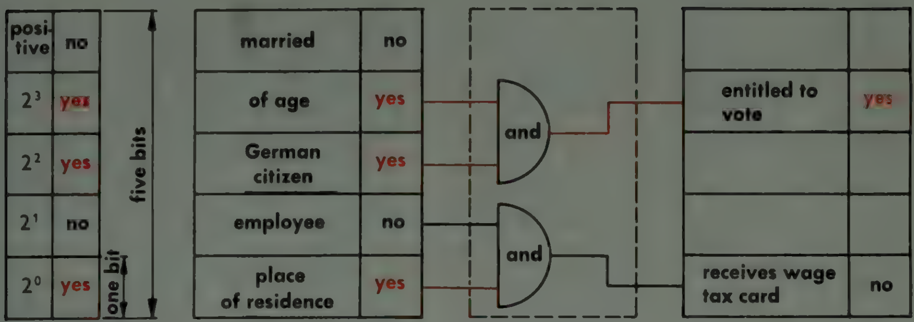


DATA PROCESSING: PRINCIPLES

The term "data" in the mathematical and technical sense denotes any facts or information, particularly as taken in, operated on, or put out by a computer or other machine for handling information. Inferences can be drawn from data. Thus the information "A is of age" and "A is a German citizen" leads to the inference "A is entitled to vote", this being so by virtue of the German election law. This selection of significant information from a set of given data is known as "data processing". In a more general sense this refers to all the operations performed on data according to prescribed plans. These operations range from the collection of raw data to the reporting of the results of calculations involving the data. If a large number of data have to be processed according to the same rules, this can be done by machines. For a limited range of subjects, any description can be replaced by a number of questions which are answered by "yes" or "no" (binary number system). With the aid of a code of this kind it is possible to represent and store data. To this end, it is necessary to have components with two readily distinguishable conditions. Depending on the speed with which the "yes" and "no" conditions are recorded and subsequently retrieved, various devices are used for the purpose: punched cards, paper tape, magnetic tape, magnetic cores, magnetic drums, tubes, transistors, in conjunction with electric impulses. A binary digit, i.e., a digit in the binary scale of notation, is commonly abbreviated as "bit". As a basic unit of information the bit is the information contained in a decision between "yes" and "no". In a data processing machine a number of bits, are combined into what is known as a "machine word". Each machine word is stored in a so-called "store location" or a "register"; the latter term refers more particularly to a device for storing one word at a time while it is actually being used. Data processing means making up a new machine word — in accordance with certain rules — from the contents of one or more registers. All conceivable rules for data processing can be built up from a limited number of simple connections such as, for example, "and", "or" and "negation". In the above example, the right to vote can be ascertained from a punched card index of the population by means of a simple "and" circuit. At the output this yields a "yes" only if there is a "yes" at both inputs. A circuit which adds the contents of a particular "storage location" as a binary number to the contents of a register is much more complex, but can nevertheless be built up from simple connective circuits. This principle is not a rational one for a purely mechanical data processing machine. It is used only for machines which operate with relays, electronic tubes or transistors. Circuits for the most frequently occurring arithmetical operations are pre-installed in the machine. Complex data processing operations (e.g., long numerical computations) can be composed from these "basic instructions". A programme-controlled machine carries out instructions in the sequence in which these occur in consecutive storage locations of a storage (a device for receiving and holding information) or a punched tape.

(Continued)

splitting-up of various items of information
into elementary statements

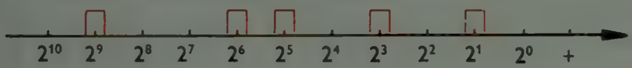


the number 13 represented in the binary system

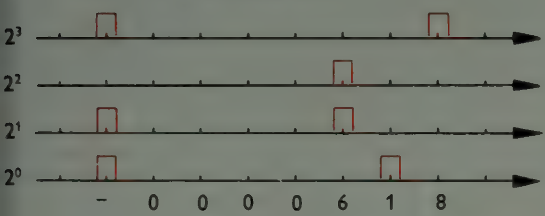
entry in card index of inhabitants

data processing

the number —618 represented by a pulse train



representation of the number —618
by four parallel pulse trains
the "+" sign is coded as "12",
the "—" sign is coded as "11"



elementary logical connections and the
usual symbols employed for them

A	B	C
no	no	no
yes	no	no
no	yes	no
yes	yes	yes

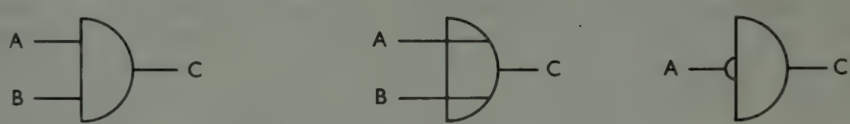
A	B	C
no	no	no
yes	no	yes
no	yes	yes
yes	yes	yes

A	C
no	yes
yes	no

and

or

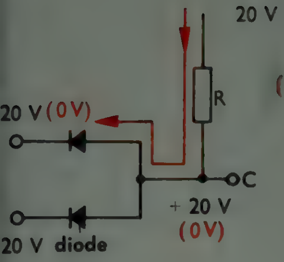
negation



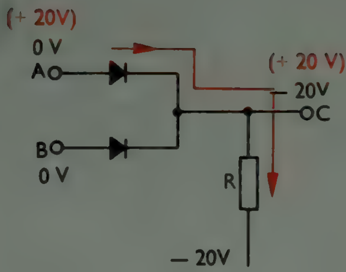
Circuits for the elementary logical connections are the elements for the building up of computer circuits. The accompanying diagram shows circuits for the case where "yes" is represented by a positive voltage pulse, "no" by a negative pulse on the absence of a pulse.

In the "and" circuit no current flows in the rest position (black); the output C therefore has a voltage of +20 V, like A and B. If at least one of the inputs A and B is given a voltage 0 V, a current will flow through the diode, and a voltage drop will occur in the resistance R. Since the diode offers practically zero resistance to this direction of flow, C acquires the voltage 0 V. In the "or" circuit both diodes block the flow of current if A and B are both at 0 V. If A or B or both these inputs are at +20 V, then C is likewise connected to +20 V, without the interposition of a resistance. In the circuit for "negation" the grid of the tube is negative in relation to the cathode. The tube then blocks the flow of current, and the time pulses proceed undisturbed via the two conductors at the anode. When a positive pulse occurs at A, the tube allows current to pass; the time pulses are suppressed or greatly diminished. The input pulses of the connective circuits are often not of a form suitable for further processing. For example, the "or" circuit gives −20 V instead of 0 V for "no"; the negation circuit does not entirely suppress the negated pulses. So-called "triggers" produce rectangular pulses of uniform height from positive pulses of arbitrary shape. In the trigger circuit illustrated, tube 2 conducts current in the rest position (black); tube 1 blocks the current, since its grid is negative in relation to the cathode; its high anode voltage gives a positive voltage (through a voltage divider) to the grid of tube 2. When a positive pulse is fed in at the input, tube 1 will begin to conduct electricity when a certain grid voltage is exceeded. Its anode voltage decreases in consequence of this, and through the voltage divider the grid voltage of tube 2 also becomes so low that this tube now blocks the flow of current. At the same time the voltage at the output rises. The transition between the two states of the circuit is therefore effected suddenly, so that positive input pulses of sufficient height will, irrespective of their shape, produce sharp rectangular pulses at the output. If the circuit is so arranged as to function symmetrically, it is called a "flip-flop" circuit. In general, a flip-flop is an electronic device or circuit with two stable states. The circuit remains in either state until the application of a signal causes it to change. It consists of two symmetrical halves, each with its own input and output terminals. The activation of one of the two halves automatically brings about the deactivation of the other half, thereby reversing the state of the flip-flop. Thus, in a vacuum-tube flip-flop, as envisaged here, when one tube is conducting, the other is cut off, and vice versa. If the digits 0 and 1 of the binary number system are associated with the two respective states, counters (totalising devices) can be constructed. The output pulse occurring at the transition from 1 to 0 is fed to the flip-flop circuit for the next higher binary place.

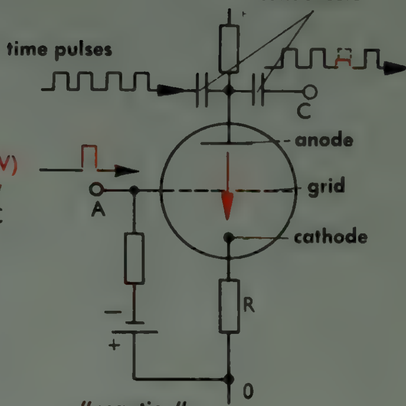
A commonly used device for memory storage is the magnetic core, which is a tiny ring made of ferrite, an easily magnetisable material. These cores are strung on insulated wires. The cores can be magnetised in either of two directions, depending on the direction of current flow in a wire. If the direction of the current is reversed, the magnetic state is changed. This bi-stable nature of the cores makes them suitable for binary representation: a core may be magnetised in one direction to store a binary 1, or in the other direction to store a binary 0. Each core stores one bit of information at a time. The cores are disposed at the intersections of a network of vertical and horizontal wires. If one-half of the current necessary to magnetise a core is sent through each of the two intersecting wires associated with any particular core, only that one core is affected. In this way the cores can be magnetised individually. The information stored in the cores is read by means of the sense wire, which passes through all the cores. When a core is switched, a pulse of current is created in the sense wire. Thus, coded information is stored by sending pulses through the appropriate cores. Reading is accomplished by detecting the effect of negative pulses on the cores.



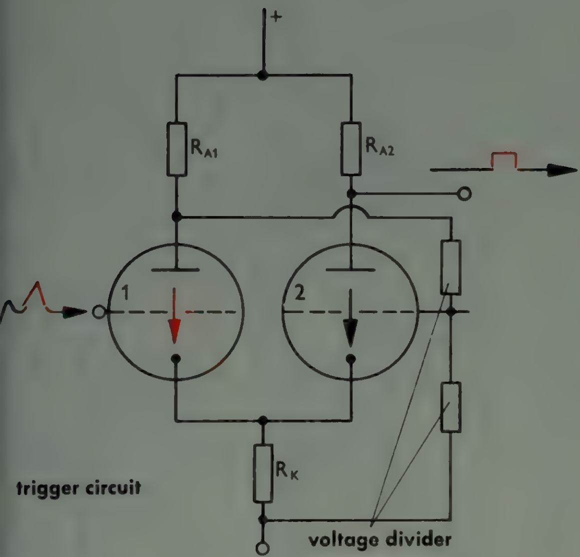
"and" circuit



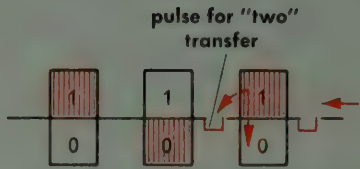
"or" circuit



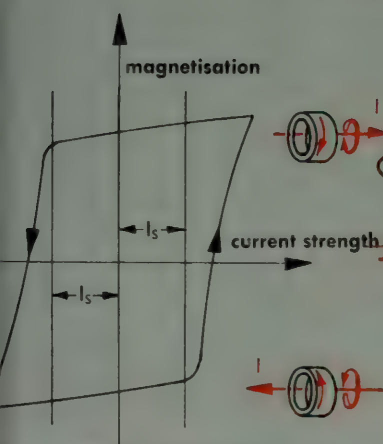
"negation"



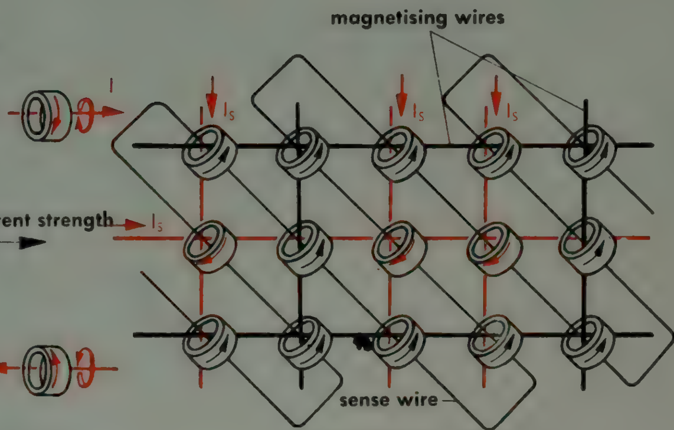
trigger circuit



counter with flip-flop circuits



agnetisation curve of a ferrite core



magnetic core plane

An International Business Machines plant technician is putting together a "mechanical brain"
Photo USIS

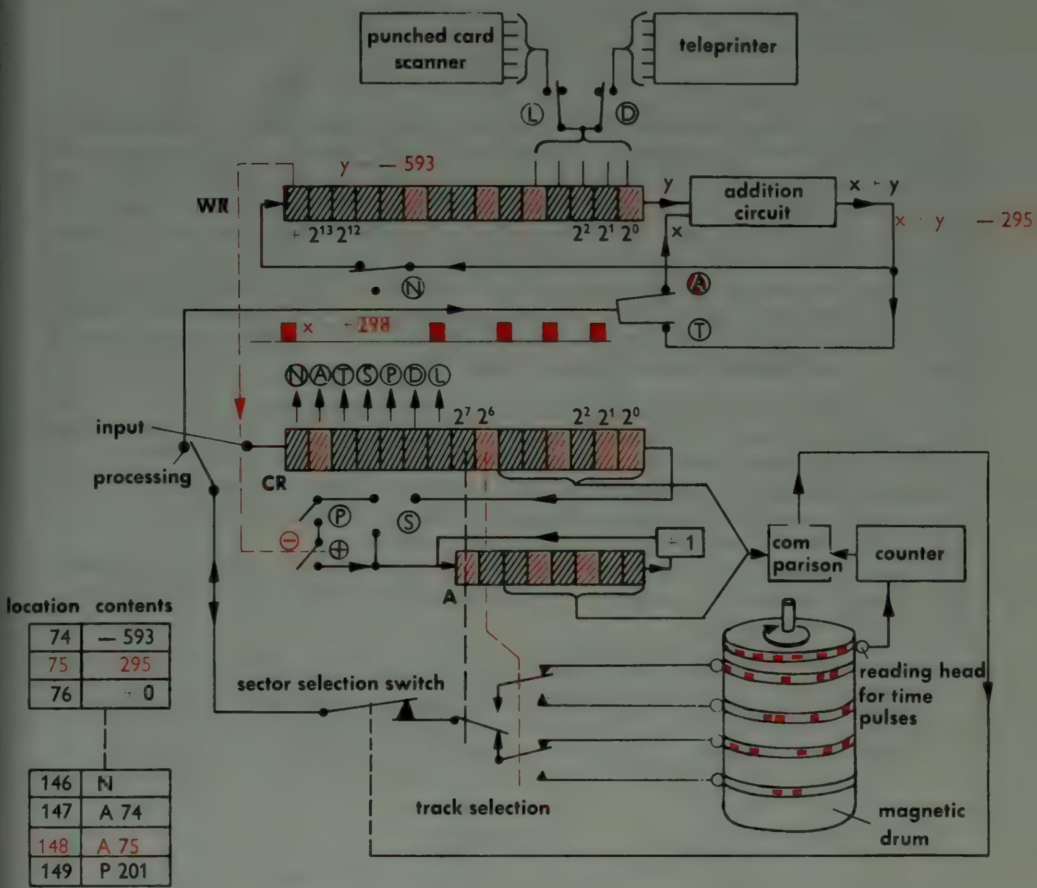
An electric computer is a universal data processing machine. Large numerical calculations are merely one of the many things it can do.

The accompanying diagram shows the principal parts of a computer, in which the individual "bits" (see page 334) of a machine word are consecutively processed as voltage pulses (serial computer). The registers consist of chains of delay elements in which a pulse is delayed for a length of time equal to the time interval between two pulses. From the end of the chain the pulse train returns to the input. When the last pulse of a word has been received, each bit can be read from the corresponding element of the chain. In the working register (WR) an addition circuit is included, through which the pulse trains coming from the contents of the working register and from the magnetic drum pass synchronously and are added together, bit by bit, starting with the lowest binary digit. The accumulator (A) accepts only eight binary digits; its contents are increased by 1 at each revolution. The computer functions in two cycles: In the first cycle the accumulator determines which word from the magnetic drum is passed to the control register (CR). Some binary digits of this instruction word, the operation part, actuate electronic switches in the pulse circuits and thus determine the operations to be performed in the next circuit. In this second cycle the eight lowest digits of the control register, the so-called address part of the instruction, determine the choice of the storage locations. In the accompanying illustration the number in the accumulator is 148. In the previous cycle the contents of location 148 was fed into the control register. This instruction contains a pulse at the position for A and closes the switch A. The address part is 75. At the instant when the location 75 reaches the read–write head of the magnetic drum, the "sector selection" switch closes for the duration of a word, and the stored pulse train, together with the hitherto existing contents of the working storage, passes through the addition circuit. Meanwhile the position of the accumulator has risen to 149; in the next cycle the instruction "P201" in location 149 is fed into the control register. The following instructions are additionally embodied in the accompanying schematic diagram:

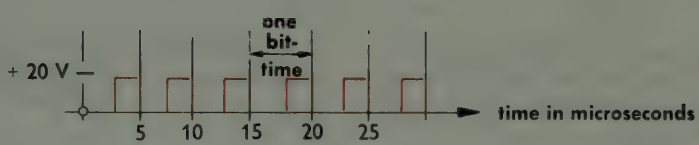
N: Resetting of the working register. The cycle in the working register is interrupted for the duration of a word.
Tn: The pulse train cycling in the working register is written in location n.
Sn: Jump of the computer operation to location n.
Pn: Jump to location n, if the contents of the working register are positive.
D: Printing out a teleprinter character corresponding to the five lowest digits of the working register.
L: The combination of holes in the punched tape scanner yields the five lowest digits of the working register.

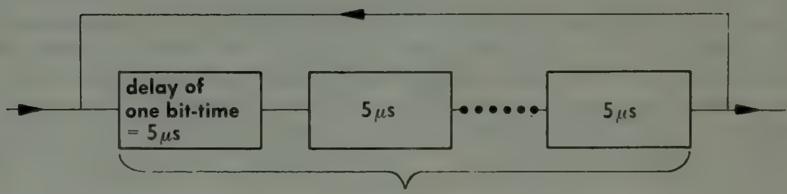
To be a practical machine, a computer requires additional instructions for shifting the contents of a register to the right or left, for reversing the algebraic sign, etc. Separate instructions for multiplication and division can, if necessary, be replaced by programmes embodying additions only. More convenient are machines in which the decimal digits 6 to 9 are represented and processed on four parallel pulse circuits (series-parallel machines).

LAYOUT OF A PROGRAMME-CONTROLLED COMPUTER



location	contents
74	— 593
75	295
76	+ 0
146	N
147	A 74
148	A 75
149	P 201





register for the storage of a pulse train with a length of 15 bits

The exponential function $y = e^x$ can be expanded into an infinite series:

$$e^x = 1 + x + \frac{x^2}{1.2} + \frac{x^3}{1.2.3} \ldots + \frac{x^k}{1.2\ldots k} \quad \ldots$$

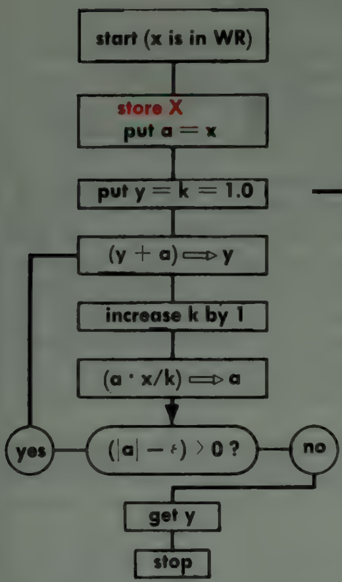
As the terms of the series steadily diminish in value, the summation can be terminated when the last term calculated is smaller than the number E which prescribes the accuracy required. We shall consider a programme for an electronic computer with decimal notation on four parallel pulse circuits. A machine word will contain 12 decimal digits and the sign digit (designation of algebraic sign, i.e., plus or minus). Decimal fractions such as, for example, 0.00347 will be written 0.347×10^{-2}, and for the first ten places behind the decimal point we shall use ten digits of the machine word. The exponent of the factor 10 (in this example: -2), increased by 50, is called the characteristic and occupies the last two places. The machine must be able to carry out the following instructions (cf. page 340):

Bn: transfer contents of storage location n to working register
Tn: transfer contents of working register to location n
+n: add contents of location n to contents of working register
×n: multiply contents of working register by contents of location n
:n: divide contents of working register by contents of location n
V: make sign of contents of working register positive
Pn: transfer address n to accumulator if contents of working register are positive
Rn: transfer address in location n to accumulator
Z: stop programme

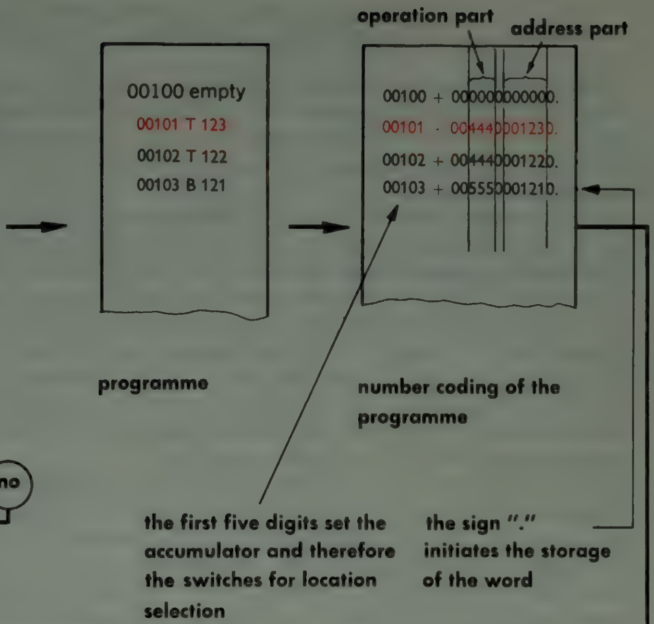
The flow diagram yields the following programme for calculating e^x:

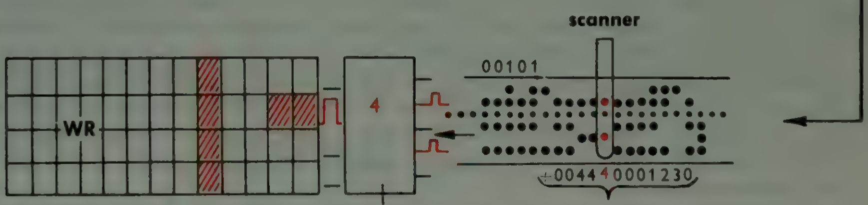
Location	Contents	Location	Contents	Location	Contents
00100	empty	00110	+121	00120	Z (R 100)
1	T 123	1	T 125	1	1.0
2	T 122	2	B 122	2	a
3	B 121	3	×123	3	X
4	T 124	4	: 125	4	Y
5	T 125	5	T 122	5	K
6	B 124	6	V	6	0.00001 = e
7	+122	7	−126	7	empty
8	T 124	8	P 1106		
9	B 125	9	B 124		

It is here assumed that, to begin with, the independent variable x is in the working register. At the stop the result is in the working register. This programme must be written on a punched tape in the number code that the machine can utilise (see accompanying illustration). In the "reading-in" of the punched tape the first five digits determine the switches for selecting a storage location, in which the 13 further characters are then stored. The calculation of e^x for an individual value of x can be done as follows: input of x into the working register, setting the accumulator to 00101, starting the computer operation. The programme only acquires practical value as a so-called subroutine, i.e., a selection of a programme which is stored once and can be used over and over again during the course of the programme to accomplish a certain operation. A fresh instruction Un, if it is in the location x, must transfer the address $x + 1$ to location n and set the accumulator at n + 1. As a result of the instruction Rn the jump back to $x + 1$ is then effected. Thus, at any position of a large programme the calculation of e^x can be initiated by U100 as a separate instruction.



flow diagram for the computation of e^x

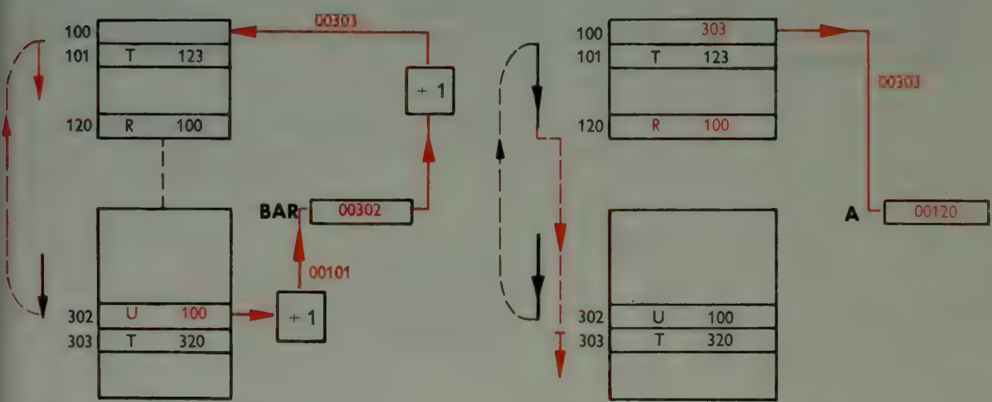




coding circuit

13 characters inserted into WR

WR; 5 characters have been read, the 6th character is being scanned and coded



effect of the instruction U

effect of the instruction R 100

subroutine jump

TRANSLATION PROGRAMME FOR A PROGRAMME-CONTROLLED COMPUTER

For translation into another language, a text must first of all be coded as a sequence of machine words. For example, one of the numbers 01 to 32 can be assigned to each of the 32 characters of a teleprinter. A word of n letters will then occupy $2n$ decimal places in the storage and may, in certain circumstances, occupy several consecutive storage locations.

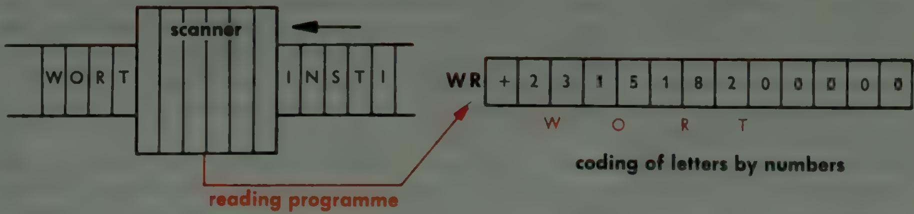
In this way it is possible to store ("memorise") a "dictionary" with entries relating to the grammatical properties of each word. The rough translation of a sentence involves comparing the words read from a punched tape with the entries in the "dictionary". If the latter contains each word in every one of the forms in which it can occur, then the process of translation will consist merely in seeking out that number in the storage which corresponds to the number read into the machine. For reasons of available storage space in the "memory" it is more advantageous only to store a list of the root words and another list containing the endings and first syllables. The root word is then sought as that number which agrees in as many digits as possible with the read-in number.

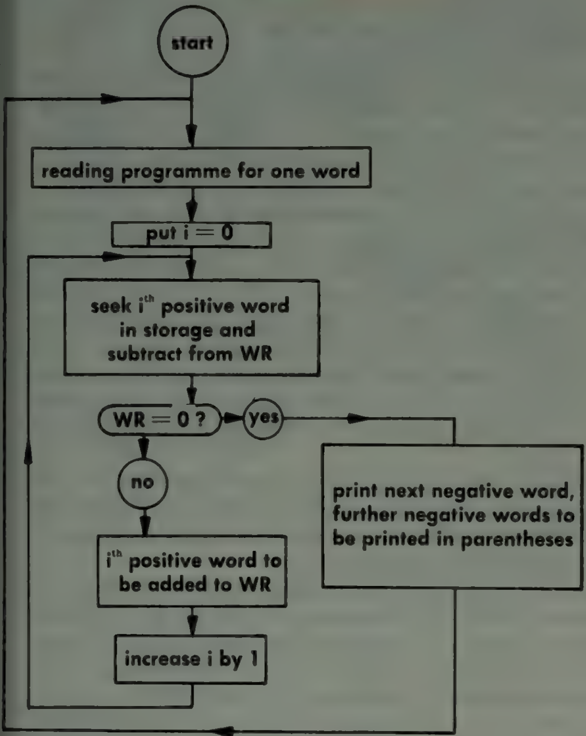
The rough translation printed out by the machine is similar to what would be obtained if each word of the text had been looked up in a dictionary, except that this laborious task has been performed by the machine. The translation then still has to be "licked into shape". However, the machine can go further than this. The "dictionary" entries stored in its "memory" can, by means of distinctive characteristics for nouns, verbs, etc., also take account of the syntax of a sentence. For example, a German sentence comprising a series of qualifying adjectives or adjectival clauses can be divided by means of parentheses, rather in the manner of a mathematical formula. In the German language the sequence of noun and reflexive pronouns corresponds to opening a parenthesis, while the verb in its active forms corresponds to closing a parenthesis. In the English translation these parentheses must be removed. This can be done by starting a new section for the storage of the translation whenever the characteristics for noun and reflexive pronoun occur in succession. When the verb occurs, the programme is again changed over to storage in the previously used storage section.

Translation programmes require a very considerable amount of storage space for "dictionaries". For this reason magnetic tapes are used as external storage media for the various "dictionaries". To obtain the information directly from the tapes would be too slow, as the tapes cannot be moved fast enough to give a sufficiently short "access time" (the time interval between the instant when information is called for from the storage and the instant when delivery is completed). For this reason those "dictionaries" which are needed at any particular stage of the translation procedure are temporarily transferred to the working storage (which may be a magnetic drum or a magnetic core storage), in which the information can be very quickly located (short access time).

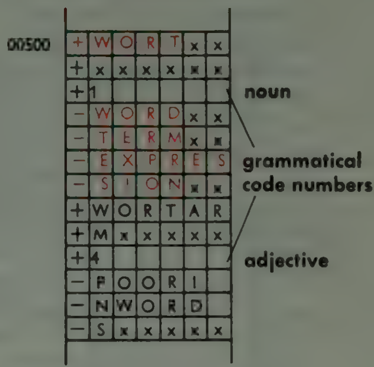
text to be translated:

DER EINTRAG, DER MIT DEM GELESENEN WORT ÜBEREINSTIMMT, WIRD GESUCHT





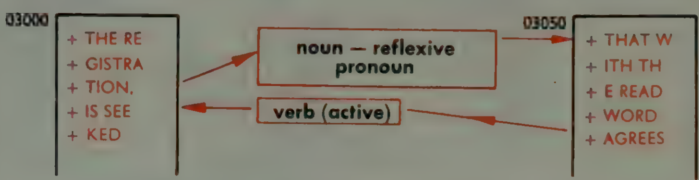
flow diagram for the production of a rough translation by computer



section of the dictionary stored in the machine

rough translation of text into English:

THE REGISTRATION, THAT WITH THE READ WORD (TERM, EXPRESSION) AGREES, IS SEEKED (SEARCHED).



sentence construction taken into account by interchanging the storage sections when certain grammatical code numbers occur

Sorting and classifying letters at main post offices is a laborious operation requiring a considerable number of staff to carry it out. This can be greatly reduced by automation. The essential requirement for an automated sorting system is that the machine shall be able to "read" the place of destination in the address on a letter. One way to achieve this is to provide the name of each town with a distinctive number written before the name. Ideally, the automatic sorting machine would have to be able to "read" the number direct from the envelope, but this is a problem that has not yet been reliably solved. Instead, the distinctive number has to be translated into characters that the machine can suitably distinguish (coding). The code elements employed for the purpose are markings applied with a special fluorescent or magnetisable paint. These are "read" in the machine, i.e., they are converted into electric pulses which are used to control mechanical selector devices in the distributing system.

The smallest unit of a letter sorting installation comprises a feed device for coded and for non-coded letters respectively, six coding stations with pre-distributors and a distributing machine (Fig. 1). Letters which have already been coded are automatically scanned and directly delivered into the channels of the distributor belt.

At the coding station (Fig. 2) the appropriate code—according to the address—is printed on the envelope. The receptacles of the intermediate stacker are swung round to the selector unit, from where they are called for by the coder. The empty receptacles swing back to the refilling position. After the code has been imprinted, the following automatic operations are performed: the letter (I) in the code printer is pushed by a carrier pin into a receptacle of the pre-distributor. The letter in the reading position (II) is brought within reach of the carrier pin by the rotation of a star rotor, which pin conveys it to the printing unit. The receptacle on the vertical conveyor opens, and another letter (III) drops into the reading position. The letter IV is passed from the selector unit to the receptacle; the selector unit then picks up another letter by suction.

The receptacles of the pre-distributor have destination stores which control the opening action of the receptacles. The latter are swung down and allow their contents to drop into one of the ten ducts (depending on the destination of the letter concerned) whence they are delivered to the appropriate receiving boxes.

From these boxes the letters are delivered to the intermediate stackers of the distributing machine (Fig. 3), in which a ring of receptacles and a ring of boxes rotate in opposite directions. The distributed letters are collected in these rotating boxes. To each receptacle is assigned an adjustable destination marker, and to each box is assigned a fixed "key" formed by magnetically controlled protective-gas contacts. The letters from the intermediate stackers or the feed device are fed individually to the synchronising device; they then pass through the code scanner and are delivered into the receptacles of the distributing machine. A code imprint in magnetic paint is scanned by means of magnetic heads; alternatively, an imprint in fluorescent paint is scanned by a device termed a photomultiplier (see page 140). The result is fed into an electronic translator which controls the magnetisation of the receptacles. When the "keys" and the boxes correspond to one another in the course of the rotation of the rings in the distribution machine, the protective-gas contacts close, the bottom of the receptacle opens, and the letter drops into the appropriate box.

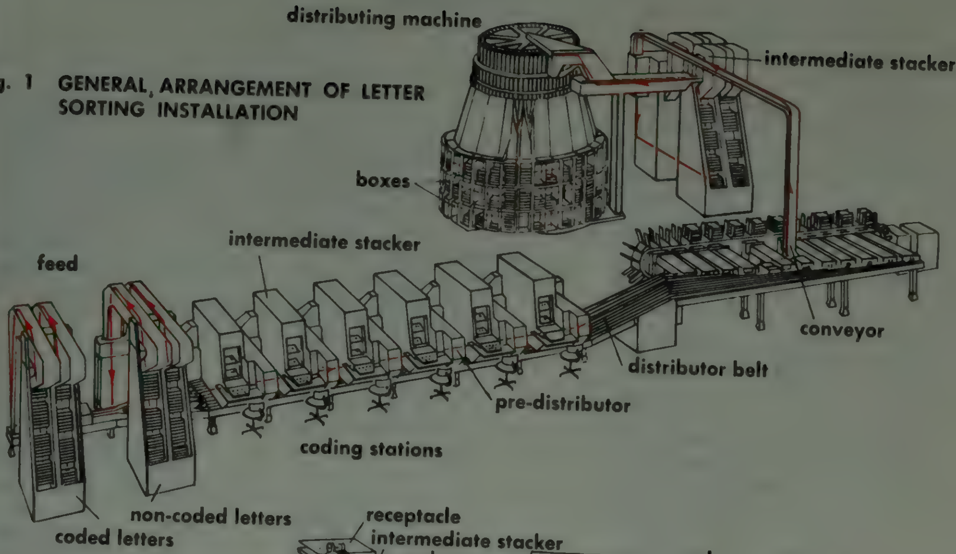


Fig. 1 GENERAL ARRANGEMENT OF LETTER SORTING INSTALLATION

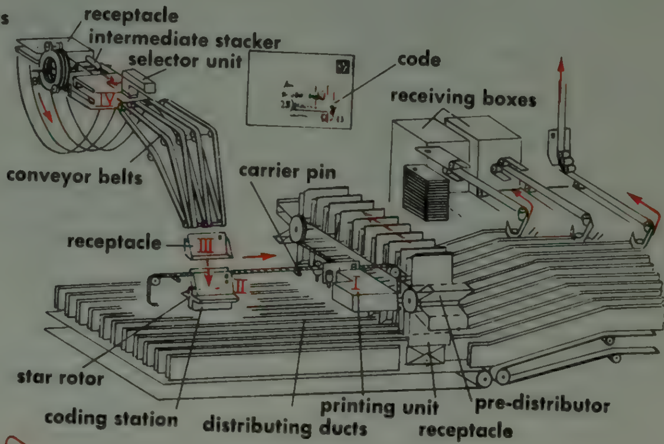


Fig. 2 PROCEDURE AT CODING STATION

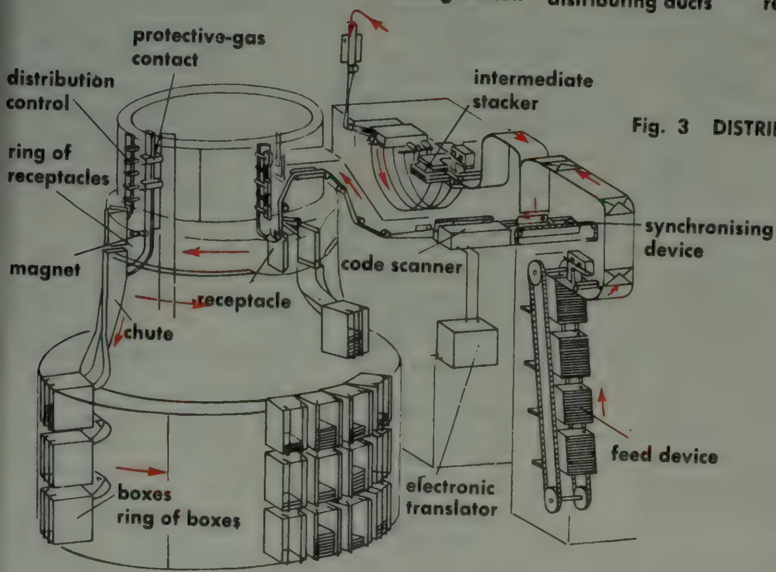


Fig. 3 DISTRIBUTING MACHINE

INDEX

This book is a production of
Edito-Service S.A., Geneva

Printed and Bound
U.S.A.